

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:11:02 ; Search time 948 Seconds
(without alignments)
2398.458 Million cell updates/sec

Title: SEQ3-THEN-SEQ4

Perfect score: 40

Sequence: 1 agagctgttggtcagacttatccagatttttggcggtag 40

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.1

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	25.2	63.0	163520	8 AL391259	Human DNA
c 2	25.2	63.0	183539	8 AC092268	Human DNA
c 3	24.4	61.0	182855	8 AC092292	Human DNA
c 4	24.4	61.0	188601	8 AC022254	Human DNA
c 5	24.2	60.5	110000	14 BX284653	Human DNA
c 6	24.2	60.5	187507	8 HSDJ137K2	Human DNA
c 7	23.2	58.0	3018	15 BT002723	Arabidops
c 8	23.2	58.0	3292	15 AY050402	Arabidops
c 9	23.2	58.0	110000	15 CR380948	Continuation (4 of
c 10	23.2	58.0	157285	14 AC099319	Continuation (4 of
c 11	23.2	58.0	194645	8 AC113932	Human DNA
c 12	23.2	58.0	201467	9 AC027298	Human DNA
c 13	23.2	58.0	204162	8 AC116992	Human DNA
c 14	23	57.5	82207	14 AC165609	Human DNA
c 15	22.8	57.0	168935	9 AC130715	Human DNA
c 16	22.8	57.0	212028	9 AL928713	Human DNA
c 17	22.6	56.5	474	10 G46487	Mouse DNA
c 18	22.6	56.5	2091	6 AX536900	Sequence

19	22.6	56.5	5403	8 HSM808431	BX648283 Homo sapi
20	22.6	56.5	40634	8 AC005305	AC005305 Homo sapi
21	22.6	56.5	94024	8 AC007368	AC007368 Homo sapi
c 22	22.6	56.5	104059	5 BX901938	BX901938 Zebrafish
c 23	22.6	56.5	156785	14 AC146517	AC146517 Homo sapi
c 24	22.6	56.5	157518	8 AC093525	AC093525 Homo sapi
c 25	22.6	56.5	205268	14 AC146518	AC146518 Homo sapi
c 26	22.6	56.5	227399	14 AC097687	AC097687 Rattus no
c 27	22.6	56.5	277521	14 AC120588	AC120588 Rattus no
c 28	22.4	56.0	58272	8 AC121338	AC121338 Homo sapi
c 29	22.4	56.0	67076	8 AY342401	AY342401 Homo sapi
c 30	22.4	56.0	103568	15 CNS07EGY	AL590462 DNA centr
c 31	22.4	56.0	110000	14 AC009727	AC009727 Homo sapi
c 32	22.4	56.0	153137	8 AC016595	AC016595 Homo sapi
c 33	22.4	56.0	321294	15 CNS09S4S	EX088700 DNA centr
c 34	22.2	55.5	1107	6 BD163586	BD163586 Novel pol
c 35	22.2	55.5	1107	6 AX121469	AX121469 Sequence
c 36	22.2	55.5	1294	6 AX763101	AX763101 Sequence
c 37	22.2	55.5	1600	6 AR629472	AR629472 Sequence
c 38	22.2	55.5	1600	6 AX395180	AX395180 Sequence
c 39	22.2	55.5	2175	9 AF087662	AF087662 Mus muscu
c 40	22.2	55.5	2175	9 AF095640	AF095640 Mus muscu
c 41	22.2	55.5	2175	9 AF109179	AF109179 Mus muscu
c 42	22.2	55.5	2175	9 AF113951	AF113951 Mus muscu
c 43	22.2	55.5	110000	1 BA000036	Continuation (14 o
c 44	22.2	55.5	200096	14 CT009566	CT009566 Sus scrof
c 45	22.2	55.5	234444	14 AC157119	AC157119 Bos tauru

ALIGNMENTS

RESULT 1
AL391259/c

LOCUS
DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL391259 163520 bp DNA linear PRI 18-MAY-2005
Human DNA sequence from clone Rp11-469E19 on chromosome Xp11.4-21.2
Contains a chloride intracellular channel 4 (CLIC4) pseudogene the
5' end of the USP9X gene for X chromosome ubiquitin specific
protease 9 (Drosophila fat facets related) and a CpG island,
complete sequence.

AL391259

AL391259.15 GI:11322864

HTG; CLIC4; ubiquitin; USP9X.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 163520)

Whitehead, S.

Direct Submission

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegasanger.ac.uk

Clone requests: clonerequest@sanger.ac.uk

On Nov 23, 2000 this sequence version replaced gi:11322140.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX

RP11-469E19 is from the library RPl1-11.2 constructed by the group
of Pister de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: vegas@sanger.ac.uk

 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

Location/Qualifiers

1..163520

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/map="p11.4-21.2"

/clone="RP11-469E19"

/clone_lib="RPCI-11.2"

66089..66840

/locus_tag="RP11-469E19.1-001"

/pseudo

66089..66840

/locus_tag="RP11-469E19.1-001"

/note="Match: proteins: AAH51622 F26H11.5 O00299 O15247

O95833 P35526 Q29238 Q9LNU39 Q9N2G5 Q9NVF8 Q9NZAI Q9XSA7

Q9Y696 Q9Z0W7 Q9Z1Q5"

/pseudo

/codon_start=1

/product="chloride intracellular channel 4 (CLIC4)

/pseudo

join(74725..75199,112561..112814,118090..118235,

120547..120626,123815..123931,125919..126112,

129746..129861,130056..130307,130383..130521,

132381..132533,133612..133716,137459..137665,

140011..140147,142038..142171,151880..151967,

154962..155304,156572..156667,157097..157308,

159085..159325,159560..159709,160928..161048,

AL109797.18:12187..9798,AL109797.18:10067..10345,

AL109797.18:12187..12312,AL109797.18:13662..13787,

AL109797.18:14979..15145,AL109797.18:21921..22029,

AL109797.18:22262..22408,AL109797.18:23034..23180,

AL109797.18:24198..24420,AL109797.18:26730..26950,

AL109797.18:30973..31163,AL109797.18:36179..36352,

AL109797.18:40238..40379,AL109797.18:41569..42332,

AL109797.18:42890..43013,AL109797.18:44042..44267,

AL109797.18:44772..44901,AL109797.18:48887..49072,

AL109797.18:50412..50632,AL109797.18:50719..50807,

AL109797.18:54923..55079,AL109797.18:55237..55449,

AL109797.18:56171..56266,AL109797.18:58057..62249)

/gene="USP9X"

/locus_tag="RP5-1172N10.1-001"

join(74725..75199,112561..112814,118090..118235,

120547..120626,123815..123931,125919..126112,

129746..129861,130056..130307,130383..130521,

132381..132533,133612..133716,137459..137665,

140011..140147,142038..142171,151880..151967,

154962..155304,156572..156667,157097..157308,

159085..159325,159560..159709,160928..161048,

AL109797.18:12187..9798,AL109797.18:10067..10345,

AL109797.18:12187..12312,AL109797.18:13662..13787,

AL109797.18:14979..15145,AL109797.18:21921..22029,

AL109797.18:22262..22408,AL109797.18:23034..23180,

AL109797.18:24198..24420,AL109797.18:26730..26950,

AL109797.18:30973..31163,AL109797.18:36179..36352,

AL109797.18:40238..40379,AL109797.18:41569..42332,

AL109797.18:42890..43013,AL109797.18:44042..44267,

AL109797.18:44772..44901,AL109797.18:48887..49072,

AL109797.18:50412..50632,AL109797.18:50719..50807,

AL109797.18:54923..55079,AL109797.18:55237..55449,

AL109797.18:56171..56266,AL109797.18:58057..62249)

/gene="USP9X"

/locus_tag="RP5-1172N10.1-001"

/product="ubiquitin specific protease 9, X-linked (fat

facets-like, Drosophila)"

/note="match: ESTs: AA007667.1 AA133380.1 AF197970.1

CDS

AF197971.1 AI097937.1 AI228327.1 AI526622.1 AI761675.1
 AJ392219 AW107303.1 AW608199.1 BE079462 BE694065 BE809565
 H43281.1 N34148.1 N36287.1 R80323.1 T74502.1 W52132.1
 match: cDNAs: AB028980 AF000986 AF070645 U67874 X98296
 X98296.1 Y13618 Y13619"
 join(112719..112814,118090..118235,120547..120626,
 123815..123931,125919..126112,129746..129861,
 130056..130307,130383..130521,132381..132533,
 133612..133716,137459..137665,140011..140147,
 142038..142171,151880..151967,154962..155304,
 156572..156667,157097..157308,159085..159325,
 159560..159709,160928..161048,AL109797.18:9668..9798,
 AL109797.18:10067..10345,AL109797.18:12187..12312,
 AL109797.18:13662..13787,AL109797.18:14979..15145,
 AL109797.18:22262..22408,AL109797.18:23034..23180,
 AL109797.18:24198..24420,AL109797.18:26730..26950,
 AL109797.18:30973..31163,AL109797.18:36179..36352,
 AL109797.18:40238..40379,AL109797.18:41569..42332,
 AL109797.18:42890..43013,AL109797.18:44042..44267,
 AL109797.18:44772..44901,AL109797.18:48887..49072,
 AL109797.18:50412..50632,AL109797.18:50719..50807,
 AL109797.18:54923..55079,AL109797.18:55237..55449,
 AL109797.18:58057..58194)
 /gene="USP9X"
 /locus_tag="RP5-1172N10.1-001"
 /standard_name="OTTHUMP0000023142"
 /note="match: proteins: O00507 P70398 Q8C775 Q8W73 Q8WX12
 Q93008"
 /codon_start=1
 /product="ubiquitin specific protease 9, X-linked (fat
 facets-like, Drosophila)"
 /protein_id="CAD13527.2"
 /db_xref="GI:57209221"
 /db_xref="GOA:Q8WX12"
 /db_xref="InterPro:IPR001394"
 /db_xref="UniProt/TREMBL:Q8W12"
 /translations="MTATRTGSPVGNDCQAPDQSQSPLOQNOTSPDSNSNPSP
 ATPDRQGDGAPDQDEDEPAFPHDTLAKLDIMNRPVWVPLPGELELLEAI
 DLSKGLDVKSEACQFRDGLTISTKITLDEAVSGKWEIHSVLSVELCVAKLSQM
 FPLLELLAMALNPHCFHYNGTSPCESVSSVQLPEDELFAKSPDRKPMVLDL
 NKEFTLNGFQILDRFINGSALNQVIAIALIKPFQCYEFLLTHVKYFLPIEMVP
 QLENTLDEELKEAKNEAKNDALSMIKSLKNLAKSPQCEETVKNLETFLKMLR
 QLTISSTFGKNNALNENKVISSVYVYTHRGNEPEEMLTAERMAEVIQNNILSTV
 LRSLHQPVVEKLEKILREVIKEKALTLQDLONIWAQAQKHEAIKVNVDLLAKLA
 WDSPEQLDHLPCPKASWTNASKQREKLELIRLAEADDKQGMWAKVLLNWL
 HSDVPVDMDLALSARIKLDYSQDRDTKIQWIDRFIEHLRTNKKVIPALKQI
 REICSLFGEAPQLSQTSPSPHYFRHDLINQHNHALVTLVAENLATYMSNRLVA
 RDHEDYDPTQVRLGSRYSYHQVQERINFLRLKQGLWLCAPQAKQIWKCLAEAV
 YLCDREACFKWYSKLMGDEPDLPDINKDFESNVQLQDPSLLTENGKCFERPFKAV
 NCREGLIVAKRRAYMDDLELIGLDVLRVVIQSNDDIASRAIDLKKEIYTNLGPRLQ
 VNQVVIHEDFIQSCFDRKASYDTLCVLDGDKDVNCARQEAQVAVRVLTVLRVINE
 CDSYIHEERTILPMSRAFRGKHLDSFVVRFPNGRQVDLLEVWSHNTDITGSRVCIILN
 RIKANYAHTKIELFVGGELDPADRLIQGLNKLKSLTAKITLTISSNMPSPDS
 DGASGPGNMGHNDSTGPNFVYSCLPVINGSLSPRYISFLWQADLGSNNPPLR
 DGRVLMKMPDSTTTEKRAICLDHAKLGESSLSPSLDSLFFGPGASOVLYLTVV
 YALLMPAGAFLADSDSFQFLKSGSLPVLMLTRNPNLPMADMETRGAYLNALK
 IAKLLIATIGYCHRAVAEACQGVGEVNMVQINQVTHDOAVLQALQINPESSE
 CMUNSVSLAQQISDEASRMPDIQVIRAIKLIWASGGSQVLPSPNEETIKYIE
 KTWAGNEPDLDEQVCEALEVMTLCFALLPTALDLSKKEKQWTFIIDLHCHSKT
 VRQVQEQFFLMCTRCMGRPLLTFTLLFTVLGSTARAKHSGDYFTLLHLLNY
 AYNSINPNVAEVLNNEIDWLKRIQIDVKTGTGTIETILEGLHGVKELAFQTS
 EKFPKICGCKGGANLKELDFFIPASNVLYQVRMNGELPAQAIQVCCGSPPTINAG
 FELLVALAGCVENLQIVDSLTEMYYIGTATITCEALTEWEVLPVPGPRPKGFVGL
 KNAGATCMNSVYQQLITMYSIRNGILAIISGTSDDVDDMSGEEKQDNESNVRDDY
 FGYPQPFEDKPKALSTDRKYNIGVLRHLQVIFGLHAASRLQIYVPRGFVKPLRWG
 EFPYLRQHDHDALEFNSLDLSEALKALHPAMLSKVLGSGFADQICGCPHRYPC
 EESPTLNTVDIRNHQDLSLEQVVKGLLEGANVHCEKCNKK27VKRLLIKPLRE
 VLAIOLRKPDYDWERECATKFNQYFPRPRELDMPEYTVAGVAKLGNVNPESOLIQ
 SEQSESETAGSTKYVLGVLSQASGGHYYSYIQRNGSDGERNEWYKFDGDVTE
 CRMDDEEMKNQCFGEYMGVEVDHMKMSYRQRKWNAYILFPERMTIDQDDLE
 IRYISELAITTRPHQIIMPISAIERSFKQNVQFMHNRMQYSMEYFQMKLLTNCVGY
 LNPPPGQDHLLEPEEITWISIQIAARVFLFTTGTGFTTKVVRGSASDWDYALCILRHS

```

KNVRFPMHNVFNSEYLLCEPSEVRGAFKLVIFIAHFSIQDPCPSPEAS
PGSSQAYDNLSJDLHLLRAVLLNRESEHGHQYIYFNLFMTANLGVAKTKQL
KLSVPATFMLVSLDGBGPPKIQYAEKGLSVYSQLIRCCNVSRMOSSINGNPL
PNPFGPNLSQPIQPNQVADILFVRTSYVKIIECSNSEETVKLLRPFCCWENPQF
SSTVLSBLAQVAYSITYELRPYLDLILQILLIEDSQWQTHRIHNAKLGIPDDRDGLFD
TIQSKNHYOKRAYOCIKCMVALFNSCPVAYQILOQNGDLKRWKWTWAVEWLGDELERR
PYTGNQYTYNNSPVQSNETSNGFLEBSHARWTLAKACELCPSEEPDDQDAPDE
HESPPEDALPYPHSPGSIQQNNHVGQYITGPAAHMNPORTCQAQENYEGSEE
VSPQKQKQ
join(156235..156667,157097..157306)
/genes="USP9X"
/locus_tag="RP5-1172N10.1-002"
join(156235..156667,157097..157306)
/genes="USP9X"
/locus_tag="RP5-1172N10.1-002"

```

```

Query Match      63.0%; Score 25.2; DB 8; Length 163520;
Best Local Similarity 90.0%; Pred. No. 23;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GAGCTGCTTGGTCAGACTTATCCAGATTCT 31
|||||
Db 20183 GAGCTGCTTGGTCAGACTTCTCCACATCT 20154
|||||

```

```

RESULT 2
AC092268/c
LOCUS      AC092268      183539 bp      DNA      linear      PRI 28-JUL-2004
DEFINITION Homo sapiens chromosome X clone RP11-185017 map p11.4, complete
sequence.
ACCESSION AC092268
VERSION   AC092268.4 GI:50726746
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 183539)
AUTHORS   Wen,G., Ramser,J., Blechschmidt,K., Meindl,A. and Platzer,M.
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 183539)
AUTHORS   Wen,G. and Platzer,M.
JOURNAL   Direct Submission
TITLE     Submitted (03-JUL-2001) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE 3 (bases 1 to 183539)
AUTHORS   Wen,G. and Platzer,M.
JOURNAL   Direct Submission
TITLE     Submitted (26-JAN-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE 4 (bases 1 to 183539)
AUTHORS   Wen,G. and Platzer,M.
JOURNAL   Direct Submission
TITLE     Submitted (05-AUG-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE 5 (bases 1 to 183539)
AUTHORS   Lagemann,D. and Platzer,M.
JOURNAL   Direct Submission
TITLE     Submitted (28-JUL-2004) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
COMMENT   On Jul 28, 2004 this sequence version replaced gi:22095403.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: X144
Center clone name: RP11-185017
----- Summary Statistics
Sequencing vector: M13mp18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads

```

Assembly program: Phrap; version 0.990329
 Consensus quality: 183456 bases at least Q40
 Consensus quality: 183530 bases at least Q30
 Consensus quality: 183539 bases at least Q20
 Quality coverage: 10.93x

 This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one subclone.

 Neighboring sequence information:
 This clone is overlapped by RP11-77G22.

 Sequence Quality Assessment:
 This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

FEATURES	Location/Qualifiers
source	1..183539 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="X" /map="p11.4" /clone="RP11-185017"
source	1..77611 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone="RP11-77G22" /notes="overlapping clone"
misc_feature	6224..6294 /notes="single stranded/single chemistry region"
misc_feature	83209..83236 /note="single stranded/single chemistry region"
misc_feature	86278..86287 /notes="single stranded/single chemistry region"
misc_feature	89410..89425 /notes="single stranded/single chemistry region"
misc_feature	90394..90417 /notes="pcr product sequence only"
misc_feature	93609..93618 /notes="single stranded/single chemistry region"
misc_feature	100862..101479 /notes="single stranded/single chemistry region"
misc_feature	101954..102136 /notes="single stranded/single chemistry region"
misc_feature	104000..104002 /notes="low quality region"
misc_feature	104104..104142 /notes="single clone coverage"
misc_feature	104905..104938 /notes="single stranded/single chemistry region"
misc_feature	106581..106634 /notes="single stranded/single chemistry region"
misc_feature	111844..111882 /notes="single stranded/single chemistry region"
misc_feature	117448..117452 /notes="single stranded/single chemistry region"
misc_feature	128792..128865 /notes="single stranded/single chemistry region"
misc_feature	130343..130537 /notes="single stranded/single chemistry region"
misc_feature	130485..131547 /notes="pcr product sequence only"

```

misc_feature 131220
/notes="low quality region"
misc_feature 131258..131259
/notes="low quality region"
misc_feature 131341..131343
/notes="low quality region"
misc_feature 131695..132029
/notes="single stranded/single chemistry region"
misc_feature 134563..134730
/notes="pcr product sequence only"
misc_feature 134880..134990
/notes="single stranded/single chemistry region"
misc_feature 148458..148492
/notes="single stranded/single chemistry region"
misc_feature 149777..149830
/notes="pcr product sequence only"
misc_feature 150033..150066
/notes="single clone coverage"
misc_feature 150339..150440
/notes="single stranded/single chemistry region"
misc_feature 156678
/notes="low quality region"
unsure 156678
ORIGIN
Query Match 63.0%; Score 25.2; DB 8; Length 183539;
Best Local Similarity 90.0%; Pred. No. 23;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GAGCTGCTGGTCAGACTTATCCAGATTCCT 31
|||||
Db 174548 GAGCTGCTGGTCAGACTTCTCCACAATCT 174519

RESULT 3
AC009292 182855 bp DNA linear PRI 14-MAY-2002
LOCUS Homo sapiens chromosome 15, clone RP11-34F13, complete sequence.
DEFINITION AC009292
ACCESSION AC009292.11 GI:20564447
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 182855)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,K., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
Funke,B., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Liu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tesfaye,S., Toruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (13-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 182855)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Olivier,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (05-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 182855)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Olivier,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (14-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 14, 2002 this sequence version replaced gi:20043147.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1259
Center clone name: 34_P_13
-----
Location/Qualifiers
1..182855
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/clone="RP11-34F13"
/repeat_lib="RPCT-11 Human Male BAC"
repeat_region 17..1452
/rpt_family="L1MEC"
repeat_region 1479..1694

```

```

repeat_region /rpt_family="MER58A"
1695. .1960
/rpt_family="L1MEC"
repeat_region 2007. .2320
/rpt_family="AluSx"
2678. .2770
/rpt_family="L1ME1"
3167. .3279
/rpt_family="L2"
repeat_region complement(3452. .4098)
/rpt_family="L2"
repeat_region complement(4319. .4528)
/rpt_family="MIR"
4785. .4903
/rpt_family="FLAM_A"
4904. .4935
/rpt_family="(CATA)n"
repeat_region complement(5583. .5805)
/rpt_family="MER30"
repeat_region complement(5928. .6047)
/rpt_family="L2"
repeat_region complement(6215. .6510)
/rpt_family="AluSx"
repeat_region complement(6831. .6955)
/rpt_family="L2"
6997. .7070
/rpt_family="MIR"
repeat_region 10168. .10945
/rpt_family="L1PA3"
11966. .12156
/rpt_family="MER58A"
repeat_region complement(12203. .12247)
/rpt_family="L2"
12372. .12677
/rpt_family="AluSp"
repeat_region complement(12933. .13167)
/rpt_family="MIR"
13196. .13277
/rpt_family="L2"
repeat_region complement(13612. .13725)
/rpt_family="MIR3"
repeat_region complement(13864. .14165)
/rpt_family="AluSx"
repeat_region complement(14523. .14712)
/rpt_family="MIR"
15492. .15596
/rpt_family="L1MED"
15645. .15665
/rpt_family="AT_rich"
16201. .16220
/rpt_family="(TTTTG)n"
repeat_region complement(17004. .17062)
/rpt_family="L2"
repeat_region complement(17335. .17494)
/rpt_family="MER5A"
17559. .17803
/rpt_family="MIR"
17838. .17996
/rpt_family="MIR"
18448. .18554
/rpt_family="L1MC4"
18647. .18820
/rpt_family="MIR"
19286. .19435
/rpt_family="MER5A"
repeat_region complement(19928. .20065)
/rpt_family="FLAM_C"
20583. .20734
/rpt_family="MIR3"
repeat_region complement(20915. .21128)
/rpt_family="L2"
21129. .21431
/rpt_family="AluSp"

Query Match 61.0%; Score 24.4; DB 8; Length 182855;
Best Local Similarity 82.4%; Pred. No. 47;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 CTGCTTGGTCAGACTTATCCAGATTCTTGGCGGT 38
|||||
Db 112933 CTGCTTGGTCAGATTATCCAGATGCTTGCAGT 112966

RESULT 4
AC022254/c 188601 bp DNA linear PRI 20-JUL-2002
LOCUS Homo sapiens chromosome 15, clone RP11-282M16, complete sequence.
DEFINITION AC022254
ACCESSION AC022254
VERSION AC022254.11 GI:21913035
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 188601)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 15, clone RP11-282M16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 188601)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Penestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galegan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 188601)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galegan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,

```

Liu, G., Maclean, C., Macdonald, P., Major, J., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 188601)

REFERENCE

AUTHORS

Birren, B., Nuebaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (20-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 20, 2002 this sequence version replaced gi:21591923.

All repeats were identified using RepeatMasker:

Smit, A.P.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: LS066

Center clone name: 282_M_16

FEATURES

source

Location/Qualifiers

1. .188601
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-282M16"
/clone_lib="RPC1-11 Human Male BAC"

repeat_region

275..351

/rpt_family="(TA)n"

complement(852..1154)

/rpt_family="AluJb"

1155..1414

/rpt_family="L1ME4A"

1417..1439

/rpt_family="(TG)n"

complement(1449..1607)

/rpt_family="L1MC4"

2063..2994

/rpt_family="HALL1"

complement(2999..3294)

/rpt_family="AluJo"

3489..3630

/rpt_family="L1R54"

repeat_region
3631..4076
/rpt_family="LOR1b"
4077..4442
/rpt_family="L1R54"
4495..5044
/rpt_family="L1MB7"
complement(5045..5207)
/rpt_family="FAM"
5208..5424
/rpt_family="L1MB7"
complement(5425..5719)
/rpt_family="AluJb"
5720..5776
/rpt_family="L1MB7"
complement(5779..6070)
/rpt_family="AluSg"
6087..6435
/rpt_family="L1MB3"
complement(7060..7479)
/rpt_family="MSTB"
7883..7949
/rpt_family="(TG)n"
complement(7951..8247)
/rpt_family="AluSx"
complement(8248..8550)
/rpt_family="AluJo"
complement(8646..9063)
/rpt_family="MLT1J1"
9236..9331
/rpt_family="FLAM A"
complement(9603..9929)
/rpt_family="MLT1C"
10093..11120
/rpt_family="L1ME4A"
11311..11549
/rpt_family="L1MB4A"
11614..11644
/rpt_family="AT_rich"
11810..11871
/rpt_family="AT_rich"
complement(11871..11901)
/rpt_family="L1MB4A"
12466..12509
/rpt_family="MIR"
12624..12742
/rpt_family="MIR3"
13960..14132
/rpt_family="L1MB7"
14133..14190
/rpt_family="L1MC2"
14191..14218
/rpt_family="L1MB7"
14219..14502
/rpt_family="AluJb"
14503..15255
/rpt_family="L1MB7"
15256..15586
/rpt_family="AluSx"
15587..16115
/rpt_family="L1MB7"
16116..16410
/rpt_family="AluJb"
16411..16637
/rpt_family="L1MB7"
16639..16799
/rpt_family="AluSg/x"
16800..17070
/rpt_family="L1MB7"
17556..17762
/rpt_family="MIR"
17772..18080
/rpt_family="AluJo"
18092..18275

```

repeat_region      /rpt_family="(GAAA)n"
                    complement(18499..18567)
repeat_region      /rpt_family="MER112"
                    complement(18568..18898)
repeat_region      /rpt_family="AluJb"
                    complement(18899..19088)
repeat_region      /rpt_family="MER112"
                    19379..19422
repeat_region      /rpt_family="(CA)n"
                    20265..20430

Query Match       61.0%; Score 24.4; DB 8; Length 188601;
Best Local Similarity 82.4%; Pred.No. 47;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 CTGCTTGGTCAGACTTATCCAGATTCTTGGCGGT 38
    ||||| ||||| ||||| ||||| ||||| |||||
Db 184185 CTGCTTGGTCAGATTATCCAGATTCTTGGCAGT 184152

RESULT 5
BX284653 0
WPCOMMENT
Sequence split into 6 fragments LOCUS BX284653 Accession BX284653
Fragment Name      Begin      End
BX284653_0         1      110000
BX284653_1        100001    210000
BX284653_2        200001    310000
BX284653_3        300001    410000
BX284653_4        400001    510000
BX284653_5        500001    520332
LOCUS BX284653      520332 bp DNA linear HTG 14-JUN-2005
DEFINITION Homo sapiens chromosome 6 clone XXyac-830B11, WORKING DRAFT
SEQUENCE, 8 unordered pieces.
ACCESSION BX284653
VERSION BX284653.17 GI:35763555
KEYWORDS HTG; HTGS_PHAS1; HTGS_ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1 (bases 1 to 520332)
          Sycamore,N.
          Direct Submission
          Submitted (13-JUN-2005) Wellcome Trust Sanger Institute, Hinxton,
          Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
          Clone requests: clonerequest@sanger.ac.uk
          On Sep 25, 2003 this sequence version replaced gi:35209214.
          ----- Genome Center
          Center: Wellcome Trust Sanger Institute
          Web site: http://www.sanger.ac.uk
          Contact: vegas@sanger.ac.uk
          ----- Project Information
          Center project name: YM830B11
          ----- Summary Statistics
          Assembly program: XGAP4; version 4.5
          Sequencing vector: M13; M7815; 7% of reads
          Sequencing vector: plasmid; L08752; 92% of reads
          Chemistry: Dye-terminator; 1% of reads
          Chemistry: Dye-terminator; 79% of reads
          Chemistry: Dye-terminator; 1% of reads
          Chemistry: Dye-terminator Big Dye; 18% of reads
          Consensus quality: 513402 bases at least Q40
          Consensus quality: 514747 bases at least Q30
          Consensus quality: 515755 bases at least Q20
          Insert size: 519632; sum-of-contigs
          Quality coverage: 11.98x in Q20 bases; sum-of-contigs
          -----
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 8 contigs. The true order of the pieces
          * is not known and their order in this sequence record is

```

```

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2085: contig of 2085 bp in length
* 2086 2185: gap of 100 bp
* 2186 5001: contig of 2816 bp in length
* 5002 5101: gap of 100 bp
* 5102 7187: contig of 2086 bp in length
* 7188 7287: gap of 100 bp
* 7288 98390: contig of 91103 bp in length
* 98391 98490: gap of 100 bp
* 98491 101009: contig of 2519 bp in length
* 101010 101109: gap of 100 bp
* 101110 103915: contig of 2806 bp in length
* 103916 104015: gap of 100 bp
* 104016 135980: contig of 31965 bp in length
* 135981 136080: gap of 100 bp
* 136081 520332: contig of 384252 bp in length.

FEATURES
             Location/Qualifiers
             1..520332
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /chromosome="6"
             /clone="XXyac-830B11"
             /clone_lib="CEPH"
             1..2085
             misc_feature
             /note="assembly_fragment:00828"
             2186..5001
             misc_feature
             /note="assembly_fragment:04673"
             5102..7187
             misc_feature
             /note="assembly_fragment:08375"
             7288..98390
             misc_feature
             /note="assembly_fragment:09028"
             98491..101009
             misc_feature
             /note="assembly_fragment:11285"
             101110..103915
             misc_feature
             /note="assembly_fragment:11482"
             104016..135980
             misc_feature
             /note="assembly_fragment:11565"
             136081..520332
             misc_feature
             /note="assembly_fragment:11639"

ORIGIN
Query Match       60.5%; Score 24.2; DB 14; Length 110000;
Best Local Similarity 78.4%; Pred.No. 60;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 GCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
    ||||| ||||| ||||| ||||| ||||| |||||
Db 23513 GTTGTCTTGTCAAAATTACCAGACTCTTGGCAGTAG 23549

RESULT 6
HSDJ137K2
LOCUS
DEFINITION
Human DNA sequence from clone Rp1-137K2 on chromosome 6q25.1-25.3
Contains part of the gene for BRG1-binding protein ELD/OSA1 (p25OR
KIAA1235) (FLJ00253 FLJ22292), a novel gene and two CpG islands,
complete sequence.
ACCESSION AL049820
VERSION AL049820.23 GI:8247261
KEYWORDS HTG; CpG island; ELD/OSA1; FLJ00253; FLJ22292; KIAA1235; p25OR.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1 (bases 1 to 187507)
          Sycamore,N.
          Direct Submission

```

JOURNAL

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
 Clone requests: clonerequest@sanger.ac.uk
 On Jun 4, 2000 this sequence version replaced gi:8018160.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr6>

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vegas@sanger.ac.uk

 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 RPI1-137K2 is from the library RPI1-1 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pCYPAC2.

FEATURES

source

Location/Qualifiers

1. .187507
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="RZPD:RPCIP704K02137"
 /db_xref="taxon:9606"
 /chromosomes="6"
 /map="q25.1-25.3"
 /clone="RPI1-137K2"
 /clone_lib="RPI1-1"

misc_feature

1
 /notes="Clone left_end: RPI-137K2"

misc_feature

27741
 /notes="Clone right_end: RPI-80E10"
 join(AL162578.13:<107133..107285,
 AL162578.13:141226..141336,91557..91800,117367..117456,
 139935..140102,155519..155846,173935..174080,
 180903..181012,181742..181900,AL591545.6:2357..2566,
 AL591545.6:5619..5823,AL591545.6:7743..8008)
 /genes="RPI1-419L10.1"

/locus_tag="RP11-419L10.1-004"
 join(AL162578.13:<107133..107285,
 AL162578.13:141226..141336,91557..91800,117367..117456,
 139935..140102,155519..155846,173935..174080,
 180903..181012,181742..181900,AL591545.6:2357..2566,
 AL591545.6:5619..5823,AL591545.6:7743..8008)
 /gene="RPI1-419L10.1"

/locus_tag="RP11-419L10.1-004"
 /notes="match: cDNAs: AK025945"
 join(AL162578.13:<107133..107285,
 AL162578.13:141226..141336,91557..91800,117367..117456,
 139935..140102,155519..155846,173935..174080,
 180903..181012,181742..181900,AL591545.6:2357..2566,
 AL591545.6:5619..5823,AL591545.6:7743..7993)
 /genes="RPI1-419L10.1"

/locus_tag="RP11-419L10.1-004"
 /standard_name="OTTHUMP0000017478"
 /codon_start=1
 /protein_id="CAI42304.1"
 /db_xref="GI:57208441"
 /db_xref="GOA:Q5JU46"
 /db_xref="InterPro:IPR001606"
 /db_xref="UniProt/TREMBL:Q5JU46"

mRNA

/locus_tag="RP11-419L10.1-001"
 /standard_name="OTTHUMP0000017476"
 /notes="match: proteins: AAG36928 AAL76077 BAB85006 O14497
 Q9NQ17 Q9ULI5 Q9UP21"

/codon_start=1
 /protein_id="CAI42308.1"
 /db_xref="GI:57208445"
 /db_xref="GOA:Q5VYC4"
 /db_xref="InterPro:IPR001606"
 /db_xref="UniProt/TREMBL:Q5VYC4"

/translation="WGRSQSGPMDPMVKRQPOLYMGSNPHSQPQQSSPYPGSGYGGP
 GPRYPIGIQGRTPGAMAGMYPQQDSDGATWKETFWLPPYQGGYSGYCCQQQQ
 FYYSQQPPPHLPQOQVLPQSQQRYQYQOQDMSQEGYTRSQPLAPCKPNHDLNL
 IQQERPSLDUGSIDDLPTGTETATLSVASASGSTRSQSQSNPAPFPFPHASPH
 LSSIPGSPSPSPVSGVSGSPISPAISPGSQMPPPSPGSSSSSSSSSSSSSSSS
 MPQERTQNRQMAQYQQTGFSNPSHPSPGQGMHAGISSQGSNNGTYPQMSQY
 GQGYNSRPPAYGVSFASVSGPFGMGI SANNMHGGISQPCGAVPLGRMP SAGMQ
 NRPFQNNSSMTSPSPSQSGPGMPPMTVNRKADAAAVMQAANSASQSGS
 FPMQSGLMADSSSPYSQPMNNSISIMNTQAPYSNAPMVNNSASVGLADMSPGE
 SKLPLPLKADGKEGTGPQSPESKKSSSSSTTTGKLTIKYVELGNPEPKLWDRLITP
 MEERSPVSSLPVAGKKLDLPLRYVCKVEIGGLAQVNKKWRRLATNLNLTGTS
 ASSLKKQYIQLFAFECKIERGEPPEPVFTGDTKKQPKLQPPSPANSGLSQGPQT

CDS

/translation="QMPPYQGGYSGYCCQQQQPYYSQQPQPHLPQOQVLPQSQQ
 QRYQOQDMSQGYTRSQPLAPCKPNHDLNL IQQERPSLDUGSIDDLPTGTET
 ATL SAVASGSTRSQSQDQSNPAQSPFSPHSLSSI PGSPSPSPVSGVSGNQSRS
 GLSPASIPGSPMPQPPQSGSSSSSHALQSMPQERGTQNRQMAQYQOQDTPGS
 MAFHPSFGQMEAGISSQQSNSTGYTGQMSQYQGYGQNYSRPPAYGVSFASVSG
 PGMGISANNMHGGISQPCGAVPLGRMP SAGMQNRPFQNNSSMTSPSPSQSGPG
 GNGPMTPTNRKADAAAVMQAANSASQSGFPMQSGLMADSSSPYSQPMNNS
 SLMTQAPYSNAPMVNNSASVGLADMSPGSKLPLKADGKEGTGPQSPESKSS
 DYSQSGISQOQPTPGNLPVPSPSFASVSGFHDGSDSISSPGMPKTPSPSPKSS
 TTGKLTIKYVELGNPEPKLWDRLYAFMEERSGVPSLSLPAVKKPLDLRLYVCVK
 EIGGLAQVNKKWRRLATNLNVTSSASLSKKQYIQLYFAFECKIERGEPPEPV
 FSTGDTKKQPKLQPPSPVQMEPRACDTGGSGRKASWSPARARAGSRHRTQASGR
 RQRHFLQKLDLSPPTQQVAFIPVSLPKVSGNIFVVKGK"
 join(AL355297.20:154278..154349,AL162578.13:107136..107285,
 AL162578.13:77374..77412,AL162578.13:107136..107285,
 AL162578.13:141226..141336,91557..91800,117367..117456,
 139935..140102,155519..155846,173935..174080,
 180903..181012,155519..5823,AL591545.6:2357..2566,
 AL591545.6:5619..5823,AL591545.6:11030..11168,
 AL591545.6:11426..11598,AL591545.6:17553..17703,
 AL591545.6:20199..20295,AL591545.6:22093..22876,
 AL591545.6:25254..25384,AL591545.6:27555..29749)
 /gene="RP11-419L10.1"
 /locus_tag="RP11-419L10.1-001"
 join(AL355297.20:154278..154349,AL162578.13:34987..35181,
 AL162578.13:77374..77412,AL162578.13:107136..107285,
 AL162578.13:141226..141336,91557..91800,117367..117456,
 139935..140102,155519..155846,173935..174080,
 180903..181012,AL591545.6:2357..2566,
 AL591545.6:5619..5823,AL591545.6:11030..11168,
 AL591545.6:11426..11598,AL591545.6:17553..17703,
 AL591545.6:20199..20295,AL591545.6:22093..22876,
 AL591545.6:25254..25384,AL591545.6:27555..29749)
 /gene="RP11-419L10.1"
 /locus_tag="RP11-419L10.1-001"
 note="match: ESTs: AA041273 A1475753 A1523557 A1807659
 AI1831934 AW024272 AW076016 AW172670 AW76071 AW505376
 BR258105 BE464333 BE623292 BE646469 BE675687 BE676432
 BF462904 BF577878 BF953138 BI000767 BI000771 BI541454
 BQ307298 BQ425418 H38677 W72941 W73250
 match: cDNAs: AF259792 AF468300"
 join(AL355297.20:154335..154349,AL162578.13:34987..35181,
 AL162578.13:77374..77412,AL162578.13:107136..107285,
 AL162578.13:141226..141336,91557..91800,117367..117456,
 139935..140102,155519..155846,173935..174080,
 180903..181012,AL591545.6:2357..2566,
 AL591545.6:5619..5823,AL591545.6:11030..11168,
 AL591545.6:11426..11598,AL591545.6:17553..17703,
 AL591545.6:20199..20295,AL591545.6:22093..22876,
 AL591545.6:25254..25384,AL591545.6:27555..29279)
 /gene="RP11-419L10.1"
 /locus_tag="RP11-419L10.1-001"

CDS

/translation="WGRSQSGPMDPMVKRQPOLYMGSNPHSQPQQSSPYPGSGYGGP
 GPRYPIGIQGRTPGAMAGMYPQQDSDGATWKETFWLPPYQGGYSGYCCQQQQ
 FYYSQQPPPHLPQOQVLPQSQQRYQYQOQDMSQEGYTRSQPLAPCKPNHDLNL
 IQQERPSLDUGSIDDLPTGTETATLSVASASGSTRSQSQSNPAPFPFPHASPH
 LSSIPGSPSPSPVSGVSGSPISPAISPGSQMPPPSPGSSSSSSSSSSSSSSSS
 MPQERTQNRQMAQYQQTGFSNPSHPSPGQGMHAGISSQGSNNGTYPQMSQY
 GQGYNSRPPAYGVSFASVSGPFGMGI SANNMHGGISQPCGAVPLGRMP SAGMQ
 NRPFQNNSSMTSPSPSQSGPGMPPMTVNRKADAAAVMQAANSASQSGS
 FPMQSGLMADSSSPYSQPMNNSISIMNTQAPYSNAPMVNNSASVGLADMSPGE
 SKLPLPLKADGKEGTGPQSPESKKSSSSSTTTGKLTIKYVELGNPEPKLWDRLITP
 MEERSPVSSLPVAGKKLDLPLRYVCKVEIGGLAQVNKKWRRLATNLNLTGTS
 ASSLKKQYIQLFAFECKIERGEPPEPVFTGDTKKQPKLQPPSPANSGLSQGPQT

gene

/translation="QMPPYQGGYSGYCCQQQQPYYSQQPQPHLPQOQVLPQSQQ
 QRYQOQDMSQGYTRSQPLAPCKPNHDLNL IQQERPSLDUGSIDDLPTGTET
 ATL SAVASGSTRSQSQDQSNPAQSPFSPHSLSSI PGSPSPSPVSGVSGNQSRS
 GLSPASIPGSPMPQPPQSGSSSSSHALQSMPQERGTQNRQMAQYQOQDTPGS
 MAFHPSFGQMEAGISSQQSNSTGYTGQMSQYQGYGQNYSRPPAYGVSFASVSG
 PGMGISANNMHGGISQPCGAVPLGRMP SAGMQNRPFQNNSSMTSPSPSQSGPG
 GNGPMTPTNRKADAAAVMQAANSASQSGFPMQSGLMADSSSPYSQPMNNS
 SLMTQAPYSNAPMVNNSASVGLADMSPGSKLPLKADGKEGTGPQSPESKSS
 DYSQSGISQOQPTPGNLPVPSPSFASVSGFHDGSDSISSPGMPKTPSPSPKSS
 TTGKLTIKYVELGNPEPKLWDRLYAFMEERSGVPSLSLPAVKKPLDLRLYVCVK
 EIGGLAQVNKKWRRLATNLNVTSSASLSKKQYIQLYFAFECKIERGEPPEPV
 FSTGDTKKQPKLQPPSPVQMEPRACDTGGSGRKASWSPARARAGSRHRTQASGR
 RQRHFLQKLDLSPPTQQVAFIPVSLPKVSGNIFVVKGK"
 join(AL355297.20:154278..154349,AL162578.13:107136..107285,
 AL162578.13:77374..77412,AL162578.13:107136..107285,
 AL162578.13:141226..141336,91557..91800,117367..117456,
 139935..140102,155519..155846,173935..174080,
 180903..181012,155519..5823,AL591545.6:2357..2566,
 AL591545.6:5619..5823,AL591545.6:11030..11168,
 AL591545.6:11426..11598,AL591545.6:17553..17703,
 AL591545.6:20199..20295,AL591545.6:22093..22876,
 AL591545.6:25254..25384,AL591545.6:27555..29749)
 /gene="RP11-419L10.1"

mRNA

/translation="WGRSQSGPMDPMVKRQPOLYMGSNPHSQPQQSSPYPGSGYGGP
 GPRYPIGIQGRTPGAMAGMYPQQDSDGATWKETFWLPPYQGGYSGYCCQQQQ
 FYYSQQPPPHLPQOQVLPQSQQRYQYQOQDMSQEGYTRSQPLAPCKPNHDLNL
 IQQERPSLDUGSIDDLPTGTETATLSVASASGSTRSQSQSNPAPFPFPHASPH
 LSSIPGSPSPSPVSGVSGSPISPAISPGSQMPPPSPGSSSSSSSSSSSSSSSS
 MPQERTQNRQMAQYQQTGFSNPSHPSPGQGMHAGISSQGSNNGTYPQMSQY
 GQGYNSRPPAYGVSFASVSGPFGMGI SANNMHGGISQPCGAVPLGRMP SAGMQ
 NRPFQNNSSMTSPSPSQSGPGMPPMTVNRKADAAAVMQAANSASQSGS
 FPMQSGLMADSSSPYSQPMNNSISIMNTQAPYSNAPMVNNSASVGLADMSPGE
 SKLPLPLKADGKEGTGPQSPESKKSSSSSTTTGKLTIKYVELGNPEPKLWDRLITP
 MEERSPVSSLPVAGKKLDLPLRYVCKVEIGGLAQVNKKWRRLATNLNLTGTS
 ASSLKKQYIQLFAFECKIERGEPPEPVFTGDTKKQPKLQPPSPANSGLSQGPQT

TITLE NISC Comparative Sequencing Initiative
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 157285)
 AUTHORS Green,E.D.
 TITLE Direct Submission
 JOURNAL Submitted (09-NOV-2001) NIH Intramural Sequencing Center, 8717
 GroveMont Circle, Gaithersburg, MD 20877, USA
 REFERENCE 3 (bases 1 to 157285)
 AUTHORS Green,E.D.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAY-2002) NIH Intramural Sequencing Center, 8717
 GroveMont Circle, Gaithersburg, MD 20877, USA
 COMMENT On May 30, 2002 this sequence version replaced gi:16874852.

----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@nih.gov
 ----- Project Information
 Center project name: cuh
 Center clone name: 552N09

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 155647 bases at least Q40
 Consensus quality: 156282 bases at least Q30
 Consensus quality: 156519 bases at least Q20
 Insert size: 152000; agarose-fp
 Inert size: 156685; sum-of-contigs
 Quality coverage: 11.71x in Q20 bases; agarose-fp
 Quality coverage: 11.36x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 13117: contig of 13117 bp in length
 * 13118 13217: gap of unknown length
 * 13218 21505: contig of 8288 bp in length
 * 21506 21605: gap of unknown length
 * 21606 36207: contig of 14602 bp in length
 * 36208 36308: gap of unknown length
 * 36308 95535: contig of 59228 bp in length
 * 95535 95636: gap of unknown length
 * 95636 115087: contig of 19452 bp in length
 * 115088 115187: gap of unknown length
 * 115188 154461: contig of 39274 bp in length
 * 154462 154561: gap of unknown length
 * 154562 157285: contig of 2724 bp in length.

FEATURES

source

1. .157285
 /organism="Felis catus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9685"

misc_feature /clone="RP86-552N9"
 1. .22574
 /note="clone overlaps with GenBank Accession Number AC099317 clone RP86-503B21 (center project name cug)"
 misc_feature 1. .13117
 /note="assembly_fragment
 clone_end:77
 vector_side:left"
 gap 13118. .13217
 /estimated_length=unknown
 misc_feature 13218. .21505
 /note="assembly_fragment"
 gap 21506. .21605
 /estimated_length=unknown
 misc_feature 21606. .36207
 /note="assembly_fragment"
 gap 36208. .36307
 /estimated_length=unknown
 misc_feature 36308. .95535
 /note="assembly_fragment"
 misc_feature 88773. .157285
 /note="clone overlaps with GenBank Accession Number AC099318 clone RP86-552L21 (center project name cui)"
 gap 95536. .95635
 /estimated_length=unknown
 misc_feature 95636. .115087
 /note="assembly_fragment"
 gap 115088. .115187
 /estimated_length=unknown
 misc_feature 115188. .154461
 /note="assembly_fragment"
 gap 154462. .154561
 /estimated_length=unknown
 misc_feature 154562. .157285
 /note="assembly_fragment
 clone_end:SP6
 vector_side:right"

ORIGIN

Query Match 58.0%; Score 23.2; DB 14; Length 157285;
 Best Local Similarity 77.8%; Pred.No.1.4e+02;
 Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AGAGCTGCTTGGTCACACTTATCCAGATTCTTGCGC 36
 ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 44654 AGAGCTGCTTACTTCACACTTCAGTCTCTTGAG 44619

RESULT 11

AC113932/c 194645 bp DNA linear PRI 01-OCT-2002
 LOCUS Homo sapiens chromosome 3 clone RP11-757H2, complete sequence.
 DEFINITION AC113932
 ACCESSION AC113932
 VERSION AC113932.2 GI:23396280
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE 1 (bases 1 to 194645)

AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
 Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and
 Haugen,E.D.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 194645)

AUTHORS Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.

TITLE Direct Submission

JOURNAL Submitted (05-MAR-2002) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA

REFERENCE 3 (bases 1 to 194645)

AUTHORS	Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z., Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and Haugen,B.D.	2922	3134	8631	8742	1424	1389
TITLE	Direct Submission	2654	2758	8807	8742	927	953
JOURNAL	Submitted (01-OCT-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA	882	883	4881	4847	1296	1277
COMMENT	On Oct 1, 2002 this sequence version replaced gi:19111737.	2297	2428	853	857	1762	1757
	----- Genome Center	7254	7313	82	<800	8246	8104
	Center: University of Washington Genome Center	309	<800	2437	2490	4727	4786
	Center Code: UWGC	2403	2428	2566	2752	199	<800
	Web site: http://www.genome.washington.edu	3079	3254	771	771	592	<800
	Contact: uwgchgs@u.washington.edu	721	725	338	<800	3698	3727
	----- Project Information	25979	26263	6143	6221	6206	6251
	Center project name: chr-3	2457	2650	10330	10255	3901	3855
	Center clone name: RP11-757H2 (bc0771)	2735	2892	8577	8742	6476	6489
	----- Summary Statistics	11178	10948	20801	21093	1003	953
	Sequencing vector: plasmid; 100% of reads	4127	4171	533	<800	1544	1515
	Chemistry: Dye-terminator ET; 55% of reads	3491	3516	2749	2752	14199	14394
	Chemistry: Dye-terminator Big Dye; 45% of reads	7816	7877	1251	1209	451	<800
	Assembly program: Phrap; version 0.990319	4231	4171	7011	7054	6235	6489
	Consensus quality: 194343 bases at least Q40	1832	1807	3029	3151	1415	1389
	Consensus quality: 194638 bases at least Q30	6822	6867	509	<800	1579	1515
	Consensus quality: 194645 bases at least Q20	5708	5684	9188	9194	10496	10366
	Insert size: 194645; sum-of-contigs	1047	1049	9864	9761	225	<800
	Quality coverage: 7.5x in Q20 bases; sum-of-contigs	1710	1674	5084	5065	2734	2759
	-----	63	<800	7869	7940	5415	5367
	Overlapping Sequences:	1232	1251	2358	2490	5415	5367
	5': RP11-41F5 (UWGC:bc0157) AC020626	5849	5921	15440	15084	1511	1515
	3': RP11-668D4 (UWGC:bc0770) AC116992, 86357-bp overlap	4169	4171	9764	9761	425	<800
	-----	224	<800	118	<800	2871	2759
	Sequence Quality Assessment:	1249	1251	643	<800	4101	4089
	This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.	282	<800	2723	2752	893	953
	All manually edited bases have been reduced to quality zero.	15810	15569	4369	4327	6764	6799
	Quality levels above 40 are expected to have less than 1 error in 10,000 bp.	3477	3516	475	<800	643	<800
	Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.	4043	4171	1227	1209	5525	5562
	-----	1298	1251	4025	4055	5333	5367
	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.	818	816	15077	15084	571	<800
	-----	1704	1674	4985	5065	2573	2759
	Sequence Validation:	3246	3393	-----	-----	6824	6799
	This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below.	4110	4171	4110	4110	12309	12162
	The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.						
	----- BglII -----						
	SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt						
	----- BcoRI -----						
	SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt						
	----- HindIII -----						
	SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt						

	13641 13311 8696 8742 1073 1042						

	2067 2045 6 <800 6382 6489						

	9264 9318 559 <800 512 <800						

	404 <800 10575 10255 449 <800						

```

----- 1621      1674      5705      5814
----- 3355      3393      1360      1389
----- 4209      4171      755       <800
----- 1174      1146      708       <800
----- 4721      4656      2612      2759
----- 1957      1912      5611      5562
----- 3929      3907      4840      4786
----- 2661      2758      5589      5562
----- 3362      3516      5950      5814
----- 1751      1674      2191      2269
----- 5593      5562
----- 13084     13140
----- 622      <800
-----

FEATURES
source      Location/Qualifiers
1. .194645
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"

Query Match      58.0%; Score 23.2; DB 8; Length 194645;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 AGCTGCTTGGTCAGATTATCCAGATCTCTGGCGGT 38
    ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157191 AGCAGCTTGGTCGTGAATTATCCACATTATTACCTGT 157156

RESULT 12
AC027298
LOCUS      201467 bp DNA linear ROD 05-DEC-2000
DEFINITION Mus musculus 7 BAC RP23-266F22 (Roswell Park Cancer Institute Mouse
BAC Library) complete sequence.
ACCESSION AC027298
VERSION AC027298.20 GI:11496331
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Mus.
1 (bases 1 to 201467)
Dederich,D., Lewis,L.R., Hume,J., Edwards,C., Harris,C.,
Addison,S., Pace,A., Williams,G., Bonnin,D., Brooks,A., Brown,J.,
Buhay,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z.,
Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S.,
Fernandez,C., Ferraguto,D., Fortum-Tansey,J., Gill,R.,
Gortell,J.H., Gunaratne,P., Haller,G., Hernandez,J.,
Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., Kelly,S.,
Kovat,C., Liu,J., Liu,M., Loulsegel,H., Lozado,R.J., Martin,R.,
Massey,E., McLeod,M.P., Mei,G., Moore,S., Morgan,M., Morris,S.,
Neal,D., Nelson,A., Nguyen,R., Nguyen,N., Ogun,M., Parish,S.,
Perez,L., Reiter,D., Say,J., Shen,H., Vasquez,L., Watlington,S.,
Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A.,
Muzny,D.M., Rives,M., Scherer,S., Sodergren,E., Weinstein,G.,
Worley,K. and Gibbs,R.
Direct Submission
Unpublished

REFERENCE
AUTHORS      2 (bases 1 to 201467)
TITLE      Worley,K.C.
JOURNAL      Direct Submission
SUBMITTED (30-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE      3 (bases 1 to 201467)
AUTHORS      Worley,K.C.
TITLE      Direct Submission
JOURNAL      Submitted (01-DEC-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE      4 (bases 1 to 201467)
AUTHORS      Worley,K.C.
TITLE      Direct Submission
JOURNAL      Submitted (05-DEC-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT      On Dec 1, 2000 this sequence version replaced gi:11321700.
Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by
Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui
Zhang.

Exon/Intron boundaries of identified genes were chosen if there
were canonical splice junctions that maintained sequence continuity
across the splice junctions.

FEATURES
source      Location/Qualifiers
1. .201467
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /chromosome="7"
   /clone="RP23-266F22"
   complement (54. .289)
   /rpt_family="MULV"
   repeat_region
   /rpt_family="MULV"
   repeat_region
   895. 1313
   /rpt_family="MTC"
   repeat_region
   1762. 1785
   /rpt_family="AT_rich"
   repeat_region
   2815. 2852
   /rpt_family="AT_rich"
   repeat_region
   complement (315. .4840)
   /rpt_family="Lx2"
   repeat_region
   4836. 4870
   /rpt_family="L1"
   repeat_region
   complement (4912. .5804)
   /rpt_family="Lx5"
   repeat_region
   5879. 6078
   /rpt_family="Lx4"
   repeat_region
   complement (6132. .7418)
   /rpt_family="Lx4"
   repeat_region
   7429. 7451
   /rpt_family="AT_rich"
   repeat_region
   7654. 8483
   /rpt_family="L1_MM"
   repeat_region
   8568. 8648
   /rpt_family="L1"
   repeat_region
   complement (19374. .19511)
   /rpt_family="Lx"
   repeat_region
   19516. 19994
   /rpt_family="L1_MM"
   repeat_region
   complement (20009. .20523)
   /rpt_family="L1F"
   repeat_region
   complement (20527. 20825)
   /rpt_family="L1VL2"

```

```
repeat_region complement(20835. .21490)
/rpt_family="Lx4"
repeat_region complement(21486. .22307)
/rpt_family="Lx4"
repeat_region 22308. .22337
/rpt_family="(TTTG)n"
repeat_region complement(24534. .24625)
/rpt_family="L1"
repeat_region complement(24623. .24743)
/rpt_family="L1"
repeat_region 24745. .26002
/rpt_family="L1_MM"
repeat_region 26091. .26259
/rpt_family="RLTR10"
repeat_region 29240. .29697
/rpt_family="L1_MM"
repeat_region complement(30087. .30773)
/rpt_family="ETnERV2"
repeat_region complement(31847. .32382)
/rpt_family="MMETn"
repeat_region complement(32387. .33027)
/rpt_family="ETnERV2"
repeat_region complement(33013. .33334)
/rpt_family="ETnERV2"
repeat_region complement(33343. .33402)
/rpt_family="ETnERV2"
repeat_region 33403. .33884
/rpt_family="LTRIS_MM"
repeat_region complement(33885. .33949)
/rpt_family="ETnERV2"
repeat_region complement(33997. .34265)
/rpt_family="ETnERV2"
repeat_region complement(34577. .35032)
/rpt_family="ETnERV2"
misc_feature /note="Region similar to Mm#S176249 vB56h02.r1 Mus
musculus cDNA: gb=AA387402"
repeat_region complement(35348. .35455)
/rpt_family="ETnERV2"
repeat_region 35456. .35642
/rpt_family="L1_MM"
repeat_region 36292. .36343
/rpt_family="(CATATA)n"
repeat_region 36378. .36460
/rpt_family="A-rich"
repeat_region complement(36461. .36757)
/rpt_family="Lx7"
repeat_region complement(36850. .38900)
/rpt_family="Lx2"
repeat_region complement(38914. .40129)
/rpt_family="Lx2"
repeat_region complement(40132. .40249)
/rpt_family="Lx2"
repeat_region 40592. .41250
/rpt_family="L1"
repeat_region 41396. .41541
/rpt_family="L1"
repeat_region 41777. .41917
/rpt_family="L1"
repeat_region complement(41944. .42321)
/rpt_family="Lx2"
repeat_region complement(42481. .43067)
/rpt_family="Lx2"
repeat_region 49133. .50140
/rpt_family="RLTR13D"
repeat_region 54195. .54661
/rpt_family="L1_MM"
repeat_region complement(56789. .58214)
/rpt_family="L1VL2"
repeat_region 58666. .58786
/rpt_family="HERVL"
repeat_region 59692. .60261
/rpt_family="MT-INTERNAL"
repeat_region 60314. .60343
/rpt_family="(TTTA)n"
```

```
repeat_region complement(60438. .62610)
/rpt_family="Lx2"
repeat_region 65129. .65223
/rpt_family="RMER20"
repeat_region 65340. .65610
/rpt_family="Lx5"
repeat_region complement(65954. .66079)
/rpt_family="RMER20"
repeat_region 66080. .66418
/rpt_family="Lx5"
repeat_region 66419. .66458
/rpt_family="(CA)n"
repeat_region 66459. .66836
/rpt_family="Lx5"
repeat_region 66895. .67131
/rpt_family="Lx7"
repeat_region 67148. .67201
/rpt_family="(TG)n"
repeat_region complement(67264. .67400)
/rpt_family="Lx"
repeat_region complement(67402. .67684)
/rpt_family="ORR1A2"
repeat_region 67794. .67965
/rpt_family="B2"
repeat_region 68188. .68440

Query Match 58.0%; Score 23.2; DB 9; Length 201467;
Best Local Similarity 77.8%; Pred.No.1.4e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
||| ||||| ||| ||| ||||| |||||
Db 115057 CTCTTGGTCACATAGGAGATTCTTGGGGGTAG 115092

RESULT 13
AC116992
LOCUS AC116992 204162 bp DNA linear PRI 01-JUN-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-668D4, complete sequence.
ACCESSION AC116992
VERSION AC116992.2 GI:21306689
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 204162)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
Direct Submission
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 204162)
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 204162)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
Direct Submission
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On Jun 1, 2002 this sequence version replaced gi:20043132.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchtgs@u.washington.edu
----- Project Information
Center project name: chr-3
```

Center clone name: RP11-668D4 (bc0770)

----- Summary Statistics -----

Sequencing vector: Plasmid; 100% of reads
 Chemistry: Dye-terminator ET; 48% of reads
 Assembly: Dye-terminator Big Dye; 52% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 203981 bases at least Q40
 Consensus quality: 204160 bases at least Q30
 Consensus quality: 204162 bases at least Q20
 Insert size: 204162; sum-of-contigs
 Quality coverage: 8.2x in Q20 bases; sum-of-contigs

----- Overlapping Sequences -----

5': RP11-89F18 (UMGC:bc0210) AC092038, 104258-bp overlap
 3': RP11-757H2 (UMGC:bc0771) AC113932

----- Sequence Quality Assessment -----

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

----- Sequence Validation -----

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

BglII				NsiI				EcoRI			
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
3017	2927	21681	22111	8696	8802						
-----	-----	-----	-----	-----	-----						
2067	2101	1823	1842	6	<800						
-----	-----	-----	-----	-----	-----						
8151	8286	2912	2880	5030	5004						
-----	-----	-----	-----	-----	-----						
317	<800	463	<800	5757	5717						
-----	-----	-----	-----	-----	-----						
64	<800	2349	2347	12366	12038						
-----	-----	-----	-----	-----	-----						
6544	6451	5483	5381	14141	13920						
-----	-----	-----	-----	-----	-----						
1359	1317	336	<800	8270	8215						
-----	-----	-----	-----	-----	-----						
3251	3219	21	<800	4241	4265						
-----	-----	-----	-----	-----	-----						
5447	5350	878	804	15430	15387						
-----	-----	-----	-----	-----	-----						
2168	2101	216	<800	4300	4265						
-----	-----	-----	-----	-----	-----						
3980	3968	1207	1245	3793	3737						
-----	-----	-----	-----	-----	-----						

2952	2927	2771	2701	36	<800
-----	-----	-----	-----	-----	-----
1596	1551	5512	5381	2233	2206
-----	-----	-----	-----	-----	-----
2092	2101	10558	10288	354	<800
-----	-----	-----	-----	-----	-----
10117	9842	449	<800	20265	20634
-----	-----	-----	-----	-----	-----
7342	7283	6679	6545	7259	7342
-----	-----	-----	-----	-----	-----
4151	4152	1660	1594	777	767
-----	-----	-----	-----	-----	-----
1373	1317	1092	1096	3421	3402
-----	-----	-----	-----	-----	-----
1912	1876	13	<800	1008	1002
-----	-----	-----	-----	-----	-----
6673	6711	1017	1015	8656	8802
-----	-----	-----	-----	-----	-----
6255	6229	1415	1397	467	<800
-----	-----	-----	-----	-----	-----
3953	3858	3045	3050	559	<800
-----	-----	-----	-----	-----	-----
964	959	2201	2347	10575	10341
-----	-----	-----	-----	-----	-----
4308	4152	2809	2880	8632	8802
-----	-----	-----	-----	-----	-----
1882	1876	2625	2701	8807	8802
-----	-----	-----	-----	-----	-----
1605	1551	95	<800	4881	4849
-----	-----	-----	-----	-----	-----
5620	5568	570	<800	853	862
-----	-----	-----	-----	-----	-----
968	959	6870	6828	82	<800
-----	-----	-----	-----	-----	-----
6809	6711	2136	2181	2437	2458
-----	-----	-----	-----	-----	-----
10741	10472	925	922	2566	2627
-----	-----	-----	-----	-----	-----
3508	3488	783	<800	771	<800
-----	-----	-----	-----	-----	-----
3320	3219	4405	4300	338	<800
-----	-----	-----	-----	-----	-----
2121	2101	556	<800	6143	6210
-----	-----	-----	-----	-----	-----
5525	5350	1418	1397	10330	10341
-----	-----	-----	-----	-----	-----
404	<800	277	<800	8577	8802
-----	-----	-----	-----	-----	-----
2922	2927	3391	3400	20804	20634
-----	-----	-----	-----	-----	-----
2654	2759	4252	4300		
-----	-----	-----	-----		
882	876	2386	2347		
-----	-----	-----	-----		
2297	2400	1095	1096		
-----	-----	-----	-----		
7255	7283	179	<800		
-----	-----	-----	-----		
309	<800	627	<800		
-----	-----	-----	-----		
2403	2400	1400	1397		
-----	-----	-----	-----		
3079	3219	677	<800		
-----	-----	-----	-----		
721	722	4367	4300		
-----	-----	-----	-----		
25979	26167	14905	14463		
-----	-----	-----	-----		
2457	2615	2006	2042		
-----	-----	-----	-----		
2735	2927	31	<800		
-----	-----	-----	-----		
11178	10892	10	<800		

```

-----
4127      4152      756      <800
-----
3491      3488      3974      3948
-----
7816      7806      2291      2347
-----
4993      5030
-----
5415      5381
-----
168      <800
-----
2798      2880
-----
4913      4882
-----
1238      1245

Query Match      58.0%; Score 23.2; DB 8; Length 204162;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 AGCTGTTGGTCAGACTTATCCAGATTCTTGGCGGT 38
   ||||| ||||| ||||| ||||| ||||| |||||
Db 155260 AGCAGCTTGGTCTGAATTATCCACATTATTACCTGT 155295

RESULT 14
AC165609/c
LOCUS
DEFINITION Bos taurus clone CH240-169N19, *** SEQUENCING IN PROGRESS ***, 25
unordered pieces.
ACCESSION AC165609
VERSION AC165609.1 GI:70912627
KEYWORDS HTG; HTGS PHASE1.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 82207)
Muzny, D., Adams, C., Agbai II, O., Allen, C., Alsbrooks, S., Archer, P.,
Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,
Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
Chacko, J., Chanrour, M., Chavez, D., Chen, A., Chen, G., Chen, R.,
Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De
Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,
Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,
Fernandez, S., Fernandez, P., Flagg, N., Forbes, L., Foster, P.,
Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D.,
Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlen, K.,
Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P.,
Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B.,
Howell, L.T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S.,
Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F.,
Leal, S., Lee, K., Lee, S., LeGall, F.I., Lemon, S., Lewis, L., Li, B.,
Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D.,
London, P., Lopez, J., Lorenshuwa, L., Lozada, R., Luk, T., Madu, R.,
Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
McClelland, H., McPherson, J., Mercadao, C., Metzker, M.,
Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E.,
Nott, A., Nwakenleh, O., Obregon, M., Ochi-Okorie, C., Odeh, E.,
Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B.,
Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J.,

```

```

Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
Rives, C., Rodriguez, F., Rojas, A., Ruiz, S.J., Sana, M., Sanders, W.,
Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y.,
Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R.,
Svatek, A., Taylor, E., Taylor, F., Thomas, N., Thorn, R., Thornton, R.,
Trejos, Z., Umani, K., Vargo, C., Verduzco, D., Villaseana, D., Virk, D.,
Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,
Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, P., Zhang, J., Zhang, L.,
Zhang, Z., Zhou, J., Weinstein, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 82207)
Worley, K.C.
Direct Submission
Submitted (17-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help.tmc.edu
----- Project Information
Center project name: FJVR
Center clone name: CH240-169N19
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 82280 bases at least Q40
Consensus quality: 85032 bases at least Q30
Consensus quality: 86680 bases at least Q20
Estimated insert size: 111753; sum-of-contigs estimation
Quality coverage: 1x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 3012: contig of 3012 bp in length
* 3013 3112: gap of unknown length
* 3113 5319: contig of 2207 bp in length
* 5320 5419: gap of unknown length
* 5420 7498: contig of 2079 bp in length
* 7499 7598: gap of unknown length
* 7599 12514: contig of 4916 bp in length
* 12515 12614: gap of unknown length
* 12615 15270: contig of 2656 bp in length
* 15271 15370: gap of unknown length
* 15371 19365: contig of 3995 bp in length
* 19366 19465: gap of unknown length
* 22356 22356: contig of 2891 bp in length
* 22357 22357: gap of unknown length
* 22357 25936: contig of 3480 bp in length
* 25937 26037: gap of unknown length
* 26037 31198: contig of 5162 bp in length
* 31199 31298: gap of unknown length
* 31299 33725: contig of 2427 bp in length
* 33726 33825: gap of unknown length
* 33826 36615: contig of 2790 bp in length
* 36616 36715: gap of unknown length
* 36716 40219: contig of 3504 bp in length
* 40220 40319: gap of unknown length
* 40320 42578: contig of 2259 bp in length
* 42579 42678: gap of unknown length
* 42679 45272: contig of 2594 bp in length

```

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT


```

* 45273 45372: gap of unknown length
* 45373 48647: contig of 3275 bp in length
* 48648 48747: gap of unknown length
* 48748 51594: contig of 2847 bp in length
* 51595 51694: gap of unknown length
* 51695 55334: contig of 3640 bp in length
* 55335 55434: gap of unknown length
* 55435 57549: contig of 2115 bp in length
* 57550 57649: gap of unknown length
* 57650 62109: contig of 4459 bp in length
* 62109 62208: gap of unknown length
* 62209 66559: contig of 4351 bp in length
* 66560 66659: gap of unknown length
* 66660 69181: contig of 2522 bp in length
* 69182 69281: gap of unknown length
* 69282 72504: contig of 3222 bp in length
* 72504 72603: gap of unknown length
* 72604 76984: contig of 4381 bp in length
* 76985 77084: gap of unknown length
* 77085 79651: contig of 2566 bp in length
* 79651 82207: gap of unknown length
* 79751 82207: contig of 2457 bp in length.

```

FEATURES

```

source
1. .82207
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-169N19"
3013. .3112
/estimated_length=unknown
5320. .5419
/estimated_length=unknown
7499. .7598
/estimated_length=unknown
12515. .12614
/estimated_length=unknown
15271. .15370
/estimated_length=unknown
19366. .19465
/estimated_length=unknown
22357. .22456
/estimated_length=unknown
25937. .26036
/estimated_length=unknown
31199. .31298
/estimated_length=unknown
33726. .33825
/estimated_length=unknown
36616. .36715
/estimated_length=unknown
40220. .40319
/estimated_length=unknown
42579. .42678
/estimated_length=unknown
45273. .45372
/estimated_length=unknown
48648. .48747
/estimated_length=unknown
51595. .51694
/estimated_length=unknown
55335. .55434
/estimated_length=unknown
57550. .57649
/estimated_length=unknown
62109. .62208
/estimated_length=unknown
66560. .66659
/estimated_length=unknown
69182. .69281
/estimated_length=unknown
72504. .72603
/estimated_length=unknown
76985. .77084

```

```

/estimated_length=unknown
79651. .79750
/estimated_length=unknown

gap
ORIGIN
Query Match 57.5%; Score 23; DB 14; Length 82207;
Best Local Similarity 83.9%; Pred.No.1.9e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGAGCTGCTTGGTCAGACTTATCCAGATTCT 31
|||||
Db 25712 AGAGCTGCTTGGTCAGATTATGCCAGAGCCT 25682
|||||

RESULT 15
AC130715 168935 bp DNA linear ROD 27-NOV-2003
LOCUS Mus musculus BAC clone RP23-460L16 from chromosome 2, complete
DEFINITION sequence.
AC130715
ACCESSION AC130715
VERSION AC130715.3 GI:33342436
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 168935)
AUTHORS Nguyen,C. and Kozlowicz,A.
TITLE The sequence of Mus musculus BAC clone RP23-460L16
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 168935)
AUTHORS Wilson,R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 168935)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 168935)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 168935)
AUTHORS Wilson,R.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 30, 2003 this sequence version replaced gi:32490736.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
-----
Center project name: M_BA0460L16
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa and Minako Tatenno in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AL928713 and AL929149.

FEATURES

source

Location/Qualifiers

1. .168935
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/map="2"

/clone="RP23-460L16"
/clone_lib="RPCI-23"

1. .560

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region 18217..18571
/rpt_family="L1"
repeat_region 18642..19007
/rpt_family="L1"
repeat_region 19436..19551
/rpt_family="ERVK"
repeat_region 19588..19966
/rpt_family="ERVK"
repeat_region 21953..22544
/rpt_family="RMER19B"
repeat_region 22545..22666
/rpt_family="B4"
repeat_region 25044..25656
/rpt_family="L1"
repeat_region 25655..25729
/rpt_family="L1"
repeat_region 25730..26055
/rpt_family="MaLR"
repeat_region 26056..26348
/rpt_family="L1"
repeat_region 26349..26493
/rpt_family="B4"
repeat_region 26744..26902
/rpt_family="MER2_type"
repeat_region 26977..27062
/rpt_family="MaLR"
repeat_region 27063..27146
/rpt_family="Alu"
repeat_region 27234..27466
/rpt_family="RMER6A"
repeat_region 27507..27818
/rpt_family="MER2_type"
repeat_region 28602..28690
/rpt_family="B2"
repeat_region 29034..29205
/rpt_family="B4"
repeat_region 29476..29555
/rpt_family="ERV1"
repeat_region 31123..31278
/rpt_family="B4"
repeat_region 31603..31929
/rpt_family="L1"
repeat_region 32844..34072
/rpt_family="L1"
repeat_region 36213..36494
/rpt_family="B4"
repeat_region 37680..37774
/rpt_family="L1"
repeat_region 38832..39288
/rpt_family="L1"
repeat_region 39295..39495
/rpt_family="L1"
repeat_region 39818..39999
/rpt_family="MaLR"
repeat_region 40284..40336
/rpt_family="MaLR"
repeat_region 40729..40877
/rpt_family="Alu"
repeat_region 41335..41461
/rpt_family="MaLR"
repeat_region 41467..41856
/rpt_family="L1"
repeat_region 42309..42609
/rpt_family="L2"
repeat_region 43220..43550
/rpt_family="L2"
repeat_region 43568..44065
/rpt_family="MaLR"
repeat_region 44127..44851
/rpt_family="MaLR"

Query Match 57.0%; Score 22.8; DB 9; Length 168935;
Best Local Similarity 79.4%; Pred. No. 2.1e+02;

Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 6 TGCTTGGTCAGACTTATCCAGATTCTTGGCGGTA 39
 |||||
 Db 70736 TGCTCGTTTGGCTTATCCATATCTTAGCGGTA 70769

Search completed: April 18, 2006, 19:42:56
 Job time : 954 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:05:32 ; Search time 219 Seconds
(without alignments)
1217.295 Million cell updates/sec

Title: SEQ3-THEN-SEQ4
Perfect score: 40
Sequence: 1 agagctgctgtgcagactatccagattcttggcggttag 40

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.1

Searched: 496997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24.2	60.5	145985	12	ADQ97164 Human can
2	23.2	58.0	881	6	ABN98755 Arabidops
3	22.6	56.5	413	12	ADQ17678 Human sof
4	22.6	56.5	2091	6	ABQ76538 C. albica
5	22.2	55.5	1107	5	Aah66350 C. glutami
6	22.2	55.5	1164	8	ACA00491 C. glutam
7	22.2	55.5	1230	14	AEB15386 C. glutam
8	22.2	55.5	1294	10	ADD13935 C. glutam
9	22.2	55.5	1600	6	AAL46365 C. glutam
10	22.2	55.5	349980	5	Aah68528 C. glutami
11	22	55.0	584	8	ABZ53923 Aspergill
12	21.6	54.0	885	9	ADA29579 DNA encod
13	21.6	54.0	100660	13	Continuation (4 of
14	21.4	53.5	1296	4	Aah46906 cDNA enco
15	21.4	53.5	1628	13	ADX51526 Plant ful
16	21.4	53.5	1842	5	AAS87541 DNA encod
17	21.4	53.5	2378	14	ADZ60316 Murine Ey
18	21.2	53.0	282	2	AAT32713 Sequence
19	21.2	53.0	334	2	AAT32716 Third Ig

C	20	21.2	53.0	924	8	ABZ51404
C	21	21	52.5	35143	11	ACN45110
C	22	21	52.5	122186	4	AAC89560
C	23	20.8	52.0	637	3	AAF13052
C	24	20.8	52.0	637	13	ADU57093
C	25	20.8	52.0	637	14	ADZ95096
C	26	20.8	52.0	912	8	ACA28075
C	27	20.8	52.0	110000	5	AAF94800
C	28	20.6	51.5	125	2	AAI11856
C	29	20.6	51.5	369	8	ACA54523
C	30	20.6	51.5	2181	2	AAI78247
C	31	20.6	51.5	2181	4	AAI65468
C	32	20.6	51.5	2360	9	ACD26033
C	33	20.6	51.5	2360	12	ADO14234
C	34	20.6	51.5	2360	13	ADO60163
C	35	20.6	51.5	2550	13	ADR85697
C	36	20.6	51.5	2564	14	AEB56454
C	37	20.6	51.5	2688	13	ADR85110
C	38	20.6	51.5	8688	13	ADR84523
C	39	20.6	51.5	110000	13	ABD32627_3
C	40	20.4	51.0	444	13	ACN62124
C	41	20.4	51.0	526	13	ACN58640
C	42	20.4	51.0	543	6	ABS72623
C	43	20.4	51.0	590	13	ACN60593
C	44	20.4	51.0	683	3	AAF13911
C	45	20.4	51.0	683	13	ADU57952

ALIGNMENTS

RESULT 1
ADQ97164
ID ADQ97164 standard; DNA; 145985 BP.
XX
AC ADQ97164;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human cancer associated sequence HD1-08-009, SEQ ID 140.
XX
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
OS Homo sapiens.
XX
PN WO2004060304-A2.
XX
PD 22-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US041389.
XX
PR 27-DEC-2002; 2002US-00330773.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2004-543781/52.
XX
FT New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
FT cancers such as leukemia and lymphoma.
XX
PS Claim 1; SEQ ID NO 140; 199pp; English.
XX
CC The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 145985 BP; 40325 A; 30740 C; 31191 G; 43391 T; 0 U; 338 Other;

```

Query Match      60.5%; Score 24.2; DB 12; Length 145985;
Best Local Similarity 78.4%; Pred. No. 17;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 GCTGCTTGGTCAGACTTATCCAGATTCCTTGGCGGTAG 40
Db 88582 GTTGCTTGTTCAAAATTACCAAGACTCTTGGCAGTAG 88618

RESULT 2
ABN98755
ID ABN98755 standard; DNA; 881 BP.
XX
AC ABN98755;
XX
DT 01-AUG-2002 (first entry)
XX
DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 523.
XX
KW Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
KW disease; crop; thale cress; tolerance factor; insect; pathogen;
KW nutrition; ds.
XX
OS Arabidopsis thaliana.
XX
PN US2002023281-A1.
XX
PD 21-FEB-2002.
XX
PF 26-JAN-2001; 2001US-00770445.
XX
PR 27-JAN-2000; 2000US-0178472P.
XX
PA (GORL/) GORLACH J.
PA (ANY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUY/) YU Y.
PA (NAME/) RAMEAKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHAW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
XX Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX
XX WPI; 2002-403163/43.
XX
XX New Arabidopsis thaliana nucleic acid for identifying homologous genes,
PT producing compositions that modulate the expression or function of its
PT encoded protein, and mapping functional regions of protein.
XX
PS Claim 1; SEQ ID NO 523; 49pp + Sequence Listing; English.
XX
XX The invention relates to an Arabidopsis thaliana nucleic acid (I)
CC comprising a sequence capable of hybridising under stringent conditions
CC to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
CC given in the specification or its fragment. A polypeptide (II) encoded by
CC (I), a transgenic plant (III) comprising an exogenous nucleic acid or a
CC genetically modified cell (IV) comprising an exogenous nucleic acid, is

useful for screening a candidate agent for its biological effect. (I) is
useful in identifying homologous or related genes, in producing
compositions that modulate the expression or function of its encoded
protein, mapping functional regions of the protein and in studying
associated physiological pathways. (I) is also useful for the genetic
manipulation of cells, particularly plant cells. (I) is also useful in
screening assays of various plant strains to determine the strains that
are best capable of withstanding a particular disease or environmental
stress. (II) and (III) are useful for screening of biologically active
agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
pathways. The screened agents are useful in improved methods of treating
crops to prevent or treat disease. (II) are also useful in screening
programs to identify agents that mimic or enhance the action of tolerance
factors. Such agents are useful in improved methods of treating crops to
enhance their tolerance to environmental stress. (I) is also useful for
enhancing or inhibiting production of a biosynthetic product in a plant.
(III) is useful for identifying other mediators that may induce
expression of proteins of interest, for establishing the extent to which
any specific insect and/or pathogen is responsible for damage to a
particular plant, for identifying other mediators that enhance or induce
tolerance to environmental stress, for identifying factors involved in
biosynthetic pathways of nutritional, commercial, or medicinal value and
for identifying productions of nutritional, commercial or medicinal
value. (IV) is useful in the study of genetic function and regulation,
for alteration of the cellular metabolism and for screening compounds
that may affect the biological function of the gene or gene products.
CC Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from USPTO
at seqdata.uspto.gov/sequence.html?DocID=999909770445
XX
SQ Sequence 881 BP; 251 A; 154 C; 206 G; 268 T; 0 U; 2 Other;
```

PT sarcoma.
XX Example 2; SEQ ID NO 495; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 413 BP; 108 A; 106 C; 72 G; 127 T; 0 U; 0 Other;
Query Match 56.5%; Score 22.6; DB 12; Length 413;
Best Local Similarity 75.7%; Pred. No. 21;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 2 GAGCTGCTGGTCAGACTTATCCAGATTCTTGGCGGT 38
DB 276 GAGCTGCTGGTCAGACTTATCCAGATTCTTGGGT 312
RESULT 4
ABQ76538
ID ABQ76538 standard; cDNA; 2091 BP.
XX
AC ABQ76538;
XX
DT 21-NOV-2002 (first entry)
XX
DE C. albicans BAX-associated cDNA fragment SEQ ID 501.
XX
KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
KW neurodegeneration; cell death; ss.
XX
OS Candida albicans.
XX
PN WO200264766-A2.
XX
PD 22-AUG-2002.
XX
PF 21-DEC-2001; 2001WO-EP015398.
XX
PR 22-DEC-2000; 2000EP-00870318.
PR 04-JAN-2001; 2001EP-00870002.
PR 09-JAN-2001; 2001EP-00870003.
XX
PA (JANC) JANSSEN PHARM NV.
XX
PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX
XX WPI; 2002-667002/71.
DR P-PSDB; ABG93272.
XX
XX New isolated nucleic acid representing a synthetic BAX-gene, useful as
PT medicament for treating, preventing and/or alleviating yeast or fungal
PT infections or proliferative disorders, or for preventing apoptosis in
PT certain diseases.
XX
XX Claim 36; Fig 2; 344pp; English.
PS
XX
XX This invention describes a novel nucleic acid representing a synthetic
CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
CC resistant yeast or fungi, identifying, or obtaining and identifying

CC Candida spp. sequences that are differentially expressed in a pathway
CC eventually leading to programmed cell death or identifying inhibitors or
CC inhibitor sequences of Bax-induced cell death. The products of the
CC invention have cytostatic, fungicide; immunosuppressive, virucide and
CC vasotropic activity and can be used in vaccines or for gene therapy. The
CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
CC antisense molecules and antibodies are useful as medicaments or in
CC preparing a medicament for treating, preventing and/or alleviating
CC diseases associated with yeast or fungi or proliferative disorders, such
CC as cancer, or for preventing apoptosis in certain diseases. The compounds
CC or polypeptides, or the genetically modified organism are useful for
CC preparing a medicament for modifying the endogenous flora of humans and
CC other mammals. The vaccine is useful for immunising against yeast or
CC fungal infections. Apoptosis-related diseases include autoimmune disease,
CC ischaemia, diseases related with viral infections or neurodegenerations.
CC This sequence represents a polynucleotide associated with the Bax gene
CC described in the disclosure of the invention
XX
SQ Sequence 2091 BP; 725 A; 384 C; 430 G; 552 T; 0 U; 0 Other;
Query Match 56.5%; Score 22.6; DB 6; Length 2091;
Best Local Similarity 75.7%; Pred. No. 30;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 4 GCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
DB 1870 GCTGCTAGACCAGTCTATATCAAGATTATTGGAGGTG 1906
RESULT 5
AAH66350/C
ID AAH66350 standard; DNA; 1107 BP.
XX
AC AAH66350;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum coding sequence fragment SEQ ID NO: 1385.
XX
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
DR P-PSDB; AAG91131.
XX
XX Novel polynucleotides derived from Corynebacterium, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
XX Claim 8; SEQ ID NO 1385; 246pp + Sequence Listing; English.
PS
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium glutamicum, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from

CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 1107 BP; 288 A; 296 C; 288 G; 235 T; 0 U; 0 Other;

Query Match 55.5%; Score 22.2; DB 5; Length 1107;
Best Local Similarity 88.9%; Pred. No. 39;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 CAGACTTATCCAGATTCTTGGCGGTAG 40
DB 704 CCGTCATATCCAGATTCTTGGCGGTAG 678

RESULT 6
ID ACA00491/c
AC ACA00491;
XX
DT 02-JUN-2003 (first entry)
XX
DE C. glutamicum derived ORF SEQ ID 482.
XX
KW Coryneform; nucleic acid array; fermentation; culture; ds.
XX
OS Corynebacterium glutamicum.
XX
PN DE10128510-A1.
XX
PD 19-DEC-2002.
XX
PF 13-JUN-2001; 2001DE-01028510.
XX
PR 13-JUN-2001; 2001DE-01028510.
XX
PA (DEGS) DEGUSSA AG.
XX

PI Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;
XX
XX WPI; 2003-279970/28.
XX
XX New nucleic acid array useful for monitoring mRNA expression of
PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
PT from Corynebacterium glutamicum.
XX
PS Claim 1; Page 221; 709pp; German.
XX

CC This invention describes a novel nucleic acid array involving
CC Corynebacterium glutamicum polynucleotides. The arrays are used to
CC analyse C. glutamicum, particularly for monitoring a fermentation process
CC to determine expression levels of C. glutamicum cellular mRNA. Such
CC monitoring particularly differentiates between expression levels of
CC different strains of C. glutamicum and allows the adjustment of different
CC culture and fermentation conditions. ACA00010-ACA02188 represent C.
CC glutamicum derived polynucleotides described in the disclosure of the
CC invention
XX
SQ Sequence 1164 BP; 302 A; 308 C; 302 G; 252 T; 0 U; 0 Other;

Query Match 55.5%; Score 22.2; DB 8; Length 1164;
Best Local Similarity 88.9%; Pred. No. 39;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 CAGACTTATCCAGATTCTTGGCGGTAG 40
DB 761 CCGTCATATCCAGATTCTTGGCGGTAG 735

RESULT 7
ID AEB15386/c
XX AEB15386 standard; DNA; 1230 BP.
AC AEB15386;
XX
DT 22-SEP-2005 (first entry)
XX
DE C glutamicum metabolic pathway regulatory protein-encoding gene SeqID301.
XX
KW metabolism; microorganism detection; microorganism identification;
KW genome; evolution; protein production; gene; ds.
XX
OS Corynebacterium glutamicum.
XX
PN US2005153402-A1.
XX
PD 14-JUL-2005.
XX
PF 06-DEC-2004; 2004US-00006098.
XX
PR 25-JUN-1999; 99US-0141031P.
PR 01-JUL-1999; 99DE-01030476.
PR 01-JUL-1999; 99US-0142690P.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 09-JUL-1999; 99DE-01032122.
PR 09-JUL-1999; 99DE-01032128.
PR 09-JUL-1999; 99DE-01032134.
PR 09-JUL-1999; 99DE-01032206.
PR 09-JUL-1999; 99DE-01032207.
PR 14-JUL-1999; 99DE-01033003.
PR 27-AUG-1999; 99US-0151251P.
PR 31-AUG-1999; 99DE-01041390.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042124.
PR 23-JUN-2000; 2000US-00602874.
XX
PA (BADI) BASF AG.
XX
PI Pompejus M, Kroger B, Schroder H, Zelder O, Haberhauer G;
XX
XX WPI; 2005-496831/50.
XX P-PSDB; AEB15387.
XX

XX An isolated metabolic pathway regulatory polypeptide from Corynebacterium
PT glutamicum, its portion, or its variant, useful for diagnosing presence
PT or activity of C. diphtheriae in subject.
XX
PS Claim 1; SEQ ID NO 301; 65pp; English.
XX

CC This invention relates to novel isolated metabolic pathway regulatory
CC polypeptides from Corynebacterium glutamicum, designated MR proteins, and
CC the DNA sequences which encode them. The invention is useful in
CC identification of C glutamicum and related organisms, mapping of genomes
CC of organisms related to C glutamicum, identification and localization of
CC C glutamicum sequences of interest, evolutionary studies, determination
CC of MR protein regions required for function, modulation of MR protein
CC activity and modulation of cellular production of desired compounds such
CC as a fine chemical. The present sequence is that of a gene which encodes
CC a novel isolated metabolic pathway regulatory polypeptide of the
CC invention from Corynebacterium glutamicum. Note: This sequence does not
CC appear in the specification but was obtained in electronic format
CC directly from USPTO.
XX
SQ Sequence 1230 BP; 323 A; 320 C; 320 G; 267 T; 0 U; 0 Other;

Query Match 55.5%; Score 22.2; DB 14; Length 1230;
Best Local Similarity 88.9%; Pred. No. 40;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 CAGACTTATCCAGATTCTTGGCGGTAG 40

DE C glutamicum coding sequence fragment SEQ ID NO: 7063.

```

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
XX
XX OS Corynebacterium glutamicum.
XX
XX PN EF1108790-A2.
XX
XX PD 20-JUN-2001.
XX
XX PF 18-DEC-2000; 2000EP-00127688.
XX
XX PR 16-DEC-1999; 99JP-00377484.
XX
XX PR 07-APR-2000; 2000JP-00159162.
XX
XX PR 03-AUG-2000; 2000JP-00280988.
XX
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX NAkagawa S, Mizoguchi H, Ando S, Hayaashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
XX PS Disclosure; SEQ ID NO 7063; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene deriving from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
XX SQ Sequence 349980 BP; 80727 A; 91049 C; 97618 G; 80586 T; 0 U; 0 Other;
Query Match 55.5%; Score 22.2; DB 5; Length 349980;
Best Local Similarity 88.9%; Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 14 CAGACTTATCCAGATTCTTGGCGGTAG 40
DB 117041 CCGTCATATCCAGATTCTTGGCGGTAG 117015

RESULT 11
ABZ53923/c
ID ABZ53923 standard; cDNA; 564 BP.
XX
XX AC ABZ53923;
XX
XX DT 28-MAR-2003 (first entry)
XX
XX DE Aspergillus oryzae polynucleotide SEQ ID NO 3036.
XX
XX KW Aspergillus oryzae; fermentation; fungus; industrial; EST;
XX expressed sequence tag; gene; ss.
XX
XX OS Aspergillus oryzae.
XX
XX PN WO200279476-A1.
XX
XX PD 10-OCT-2002.
XX
XX PF 22-MAR-2002; 2002WO-IB000890.
XX
XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
for diagnosing a bacterial disease, as components of antibacterial

```

```

XX 30-MAR-2001; 2001JP-00098371.
XX
XX (NRAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (NARE-) NAT RES INST BREWING.
PA (NORQ ) NAT FOOD RES INST MIN AGRIC.
XX
XX Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
XX
XX WPI; 2003-046817/04.
XX
XX Detection of expression of specific Aspergillus genes for monitoring the
fermentation and growth conditions of the fungus, using DNA probes.
XX
XX Claim 1; SEQ ID NO 3036; 48pp + Sequence Listing; Japanese.
XX
XX The invention relates to a polynucleotide having any of 6006 specific
CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under
CC specific culture conditions including one or more of eutrophic,
CC oligotrophic, solid, early germination, alkaline, high temperature, low
CC temperature or maltose culture or polynucleotides stringently hybridising
CC to these sequences. The polynucleotides are useful for monitoring the
CC progress of fermentation and the growth conditions of a fungus,
CC especially of Aspergillus oryzae which is widely used in industrial
CC fermentation. Also monitoring for fungal contamination. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 564 BP; 127 A; 166 C; 126 G; 145 T; 0 U; 0 Other;
Query Match 55.0%; Score 22; DB 8; Length 564;
Best Local Similarity 73.7%; Pred. No. 41;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 3 AGCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
DB 429 AGCATCTTCGTCGGATTTCATCTCTTGGCAGAAG 392

RESULT 12
ADA29579/c
ID ADA29579 standard; DNA; 885 BP.
XX
XX AC ADA29579;
XX
XX DT 20-NOV-2003 (first entry)
XX
XX DE DNA encoding Acinetobacter baumannii protein #866.
XX
XX KW ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
XX vaccine; plant biocontrol agent.
XX
XX OS Acinetobacter baumannii.
XX
XX PN US6562958-B1.
XX
XX PD 13-MAY-2003.
XX
XX PF 04-JUN-1999; 99US-00328352.
XX
XX PR 09-JUN-1998; 98US-0088701P.
XX
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX
XX PI Breton G, Bush D;
XX
XX DR WPI; 2003-576092/54.
XX
XX DR P-PSDB; ADA33705.
XX
XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
for diagnosing a bacterial disease, as components of antibacterial

```


KW protein content; gene; ss.
XX Unidentified.
OS
XX
XX
PN US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
XX (LIU/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JB, Cao Y;
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 26266; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 1628 BP; 412 A; 432 C; 438 G; 346 T; 0 U; 0 Other;
Query Match 53.5%; Score 21.4; DB 13; Length 1628;
Best Local Similarity 80.6%; Pred. No. 93;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 GAGCTGCTTGGTCAGACTTATCCAGATTCTT 32
Db 574 GAGCTGCTTCTTCTATATTCTTCCACATTCTT 544
Search completed: April 18, 2006, 19:13:04
Job time : 222 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:13:18 ; Search time 1706.5 Seconds
(without alignments)
1096.679 Million cell updates/sec

Title: SEQ3-THEN-SEQ4

Perfect score: 40

Sequence: 1 agagctgctgtcagactatccagattcttgccggttag 40

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.1

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hc:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_gss1:*

10: gb_gss2:*

11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	24.4	61.0	931	9	AQ747245 HS 5537_A
C 2	23.2	58.0	470	1	AV561271 AV561271
C 3	23.2	58.0	501	9	AZ754258 Jhc0001_M
C 4	23.2	58.0	508	1	AI993189 701495652
C 5	23.2	58.0	576	9	BZ297325 CG3358.r1
C 6	23.2	58.0	581	8	T46556 9819 Lambda
C 7	23.2	58.0	592	9	BZ293976 CG0632.r1
C 8	23.2	58.0	599	9	BZ294269 CG0802.r1
C 9	23.2	58.0	954	5	BQ649105 AGENCOURT
C 10	23	57.5	716	9	CC530433 CH240_406
C 11	23	57.5	807	9	CC500299 CH240_338
C 12	22.8	57.0	610	9	AZ959690 2M0227020
C 13	22.8	57.0	728	4	AY105678 Zea mays
C 14	22.8	57.0	771	8	DR799776 ZM_BFB002
C 15	22.8	57.0	829	8	DR798417 ZM_BFB002
C 16	22.6	56.5	413	1	AI375672 tc09d09.x
C 17	22.6	56.5	549	2	BI192993 602947236
C 18	22.6	56.5	675	2	BF037479 601461090
C 19	22.6	56.5	682	5	EX104573 BX104573
C 20	22.6	56.5	750	9	CC068351 CSU-K33r
C 21	22.6	56.5	942	5	BQ648729 AGENCOURT
C 22	22.6	56.5	945	5	BQ668572 AGENCOURT

23	22.6	56.5	976	2	BG429091
C 24	22.4	56.0	521	10	CL328874
25	22.4	56.0	806	10	CL566935
26	22.4	56.0	814	8	CX118162
27	22.2	55.5	262	5	BY357640
C 28	22.2	55.5	438	1	AA144313
C 29	22.2	55.5	520	9	AQ083572
C 30	22.2	55.5	548	2	BB433173
C 31	22.2	55.5	600	3	BI986922
C 32	22.2	55.5	613	5	BW358295
C 33	22.2	55.5	620	5	BQ828855
C 34	22.2	55.5	693	7	CNS34408
C 35	22.2	55.5	897	6	CB183965
C 36	22	55.0	284	1	AA958222
C 37	22	55.0	376	9	AZ927744
C 38	22	55.0	377	2	BE145092
C 39	22	55.0	437	1	AI066314
40	22	55.0	543	1	AA908125
41	22	55.0	649	8	DT106262
42	22	55.0	649	8	DT110751
43	22	55.0	773	5	BW013500
C 44	22	55.0	801	7	CV120865
45	22	55.0	966	9	CC142604

ALIGNMENTS

RESULT 1
LOCUS AQ747245/c 931 bp DNA linear GSS 19-JUL-1999
DEFINITION HS_5537_AI_D05_T7A_RPCI-11 Human Male BAC Library Homo sapiens
Genomic clone Plate=1113 Col=9 Row=G, genomic survey sequence.
ACCESSION AQ747245
VERSION AQ747245.1 GI:5534403
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 931)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 1113 row: G column: 9
Seq primer: T7
Class: BAC ends
High quality sequence stop: 931.
Location/Qualifiers
1. .931
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1113 Col=9 Row=G"
/sex="male"

```

/clone lib="RPCI-11 Human Male BAC Library"
/notes="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

ORIGIN
Query Match          61.0%; Score 24.4; DB 9; Length 931;
Best Local Similarity 82.4%; Pred. No. 99;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 CTGCTGGTCAGACTTATCCAGATTCTTGCGCGT 38
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 432 CTGCTGTTTCAGATTATCCACATGCTTGCCAGT 399

RESULT 2
AV561271 470 bp mRNA linear EST 23-FEB-2004
LOCUS AV561271 Arabidopsis thaliana green siliques Columbia Arabidopsis
DEFINITION thaliana cDNA clone S148e10F 3', mRNA sequence.
ACCESSION AV561271
VERSION AV561271.1 GI:8732697
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 470)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 252-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
Location/Qualifiers
source
1..470
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SQ148e10F"
/tissue_type="green siliques"
/clone_lib="Arabidopsis thaliana green siliques Columbia"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match          58.0%; Score 23.2; DB 1; Length 470;
Best Local Similarity 77.8%; Pred. No. 2.6e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 AGCTGCTGGTCAGACTTATCCAGATTCTTGCGCGT 38
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 411 AGCTGCAATGTCATATTCATGTTTCTTGAGGT 446

RESULT 3
AZ754258 501 bp DNA linear GSS 05-FEB-2001
LOCUS Jhc0001 Mouse ES genomic library Mus musculus genomic clone 337L11,
DEFINITION genomic survey sequence.
ACCESSION AZ754258
VERSION AZ754258.1 GI:12661200
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)

```

```

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 501)
AUTHORS Chai, J.H., Locke, D.P., Ohta, T., Greal, J.M. and Nicholls, R.D.
TITLE A cluster of intronless genes in the Prader-Willi syndrome region
of mouse chromosome 7C, including a novel imprinted gene, Frat3
JOURNAL Unpublished (2001)
COMMENT Contact: Nicholls, R.D.
Department of Psychiatry, and Department of Genetics
University of Pennsylvania School of Medicine
Room 530, CRB, 415 Curie Blvd, Philadelphia, PA 19104-6140, USA
Tel: 215 898 2616
Fax: 215 898 0273
Email: robertn@mail.med.upenn.edu row: 337 column: L11
Seq primer: Sp6 sequence primer
Class: BAC ends.
Location/Qualifiers
source
1..501
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/clone="337L11"
/cell_line="RW4"
/lab_host="E. coli"
/clone_lib="Mouse ES genomic library"
/notes="Vector: pBelosAC11; Site_1: HindIII; Site_2:
HindIII; The mouse ES library was created by ligating
large-size DNA fragments from RW4 cells into the HindIII
cloning site of the pBelosAC11 vector. The recombinant
vector was then electroporated into DH10B host cells. The
inserts for the library BAC clone are ~120kb and are
flanked in the vector by T7 and Sp6 RNA promoters on
either side of the HindIII cloning site. The library was
constructed by Genome Systems (now Incyte). "

ORIGIN
Query Match          58.0%; Score 23.2; DB 9; Length 501;
Best Local Similarity 77.8%; Pred. No. 2.7e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CTGCTGGTCAGACTTATCCAGATTCTTGCGCGTAG 40
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 416 CTCTGTGGTCACATAGAGATTCCTTGCGGCTAG 451

RESULT 4
AI993189 508 bp mRNA linear EST 08-SEP-1999
LOCUS 701495852 A. thaliana, Ohio State clone set Arabidopsis thaliana
DEFINITION cDNA clone 701495852, mRNA sequence.
ACCESSION AI993189
VERSION AI993189.1 GI:5840094
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 508)
AUTHORS Chen, J., Momiya, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P.,
Gorgone, G., Burns, D., Griffin, J., Mouanoutou, M., Nguyen, D.,
Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C.,
Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S.,
Nobrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and
Hanson, D.
TITLE Arabidopsis thaliana Gene Expression MicroArray
JOURNAL Unpublished (1999)
COMMENT Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte

```


JOURNAL PUBMED COMMENT	Genome Biol. 4 (2), R10 (2003) 12620120 Contact: Wong S Department of Genetics, Smurfit Institute Trinity College Dublin Dublin 2, Ireland Tel: 353 1 6082319 Fax: 353 1 6798558 Email: swong@tcd.ie Class: plasmid ends.	LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM	BQ649105 AGENCOURT_8198363 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6283618 5', mRNA sequence. BQ649105 BQ649105.1 GI:21773277 EST. Homo sapiens (human)	954 bp mRNA linear EST 15-JUL-2002	RESULT 9 BQ649105 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM	BQ649105 AGENCOURT_8198363 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6283618 5', mRNA sequence. BQ649105 BQ649105.1 GI:21773277 EST. Homo sapiens (human)	954 bp mRNA linear EST 15-JUL-2002	RESULT 9 BQ649105 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM	BQ649105 AGENCOURT_8198363 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6283618 5', mRNA sequence. BQ649105 BQ649105.1 GI:21773277 EST. Homo sapiens (human)	954 bp mRNA linear EST 15-JUL-2002
FEATURES source	Location/Qualifiers 1..592 /organism="Candida glabrata" /mol_type="genomic DNA" /strain="CBS 138" /db_xref="taxon:5478" /clone="CG0632" /clone_lib="Candida glabrata Random Genomic Library"	Query Match 58.0%; Score 23.2; DB 9; Length 592; Best Local Similarity 77.8%; Pred. No. 2.7e+02; Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	5 CTGCTTGTCAGACTTATCCAGATTCTTGGCGGTAG 40 437 CTGGATGTTGTAACCTATCTAGATTCTTTGGCGGTAG 472	599 bp DNA linear GSS 31-OCT-2002	592	592	592	592	592	592
ORIGIN										
FEATURES source	Location/Qualifiers 1..599 /organism="Candida glabrata" /mol_type="genomic DNA" /strain="CBS 138" /db_xref="taxon:5478" /clone="CG0802" /clone_lib="Candida glabrata Random Genomic Library"	Query Match 58.0%; Score 23.2; DB 9; Length 599; Best Local Similarity 77.8%; Pred. No. 2.7e+02; Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	5 CTGCTTGTCAGACTTATCCAGATTCTTGGCGGTAG 40 437 CTGGATGTTGTAACCTATCTAGATTCTTTGGCGGTAG 472	599 bp DNA linear GSS 31-OCT-2002	599	599	599	599	599	599
ORIGIN										
FEATURES source	Location/Qualifiers 1..599 /organism="Candida glabrata" /mol_type="genomic DNA" /strain="CBS 138" /db_xref="taxon:5478" /clone="CG0802" /clone_lib="Candida glabrata Random Genomic Library"	Query Match 58.0%; Score 23.2; DB 9; Length 599; Best Local Similarity 77.8%; Pred. No. 2.7e+02; Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	5 CTGCTTGTCAGACTTATCCAGATTCTTGGCGGTAG 40 437 CTGGATGTTGTAACCTATCTAGATTCTTTGGCGGTAG 472	599 bp DNA linear GSS 31-OCT-2002	599	599	599	599	599	599
ORIGIN										
FEATURES source	Location/Qualifiers 1..599 /organism="Candida glabrata" /mol_type="genomic DNA" /strain="CBS 138" /db_xref="taxon:5478" /clone="CG0802" /clone_lib="Candida glabrata Random Genomic Library"	Query Match 58.0%; Score 23.2; DB 9; Length 599; Best Local Similarity 77.8%; Pred. No. 2.7e+02; Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	5 CTGCTTGTCAGACTTATCCAGATTCTTGGCGGTAG 40 437 CTGGATGTTGTAACCTATCTAGATTCTTTGGCGGTAG 472	599 bp DNA linear GSS 31-OCT-2002	599	599	599	599	599	599
ORIGIN										
FEATURES source	Location/Qualifiers 1..599 /organism="Candida glabrata" /mol_type="genomic DNA" /strain="CBS 138" /db_xref="taxon:5478" /clone="CG0802" /clone_lib="Candida glabrata Random Genomic Library"	Query Match 58.0%; Score 23.2; DB 9; Length 599; Best Local Similarity 77.8%; Pred. No. 2.7e+02; Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	5 CTGCTTGTCAGACTTATCCAGATTCTTGGCGGTAG 40 437 CTGGATGTTGTAACCTATCTAGATTCTTTGGCGGTAG 472	599 bp DNA linear GSS 31-OCT-2002	599	599	599	599	599	599
ORIGIN										
FEATURES source	Location/Qualifiers 1..599 /organism="Candida glabrata" /mol_type="genomic DNA" /strain="CBS 138" /db_xref="taxon:5478" /clone="CG0802" /clone_lib="Candida glabrata Random Genomic Library"	Query Match 58.0%; Score 23.2; DB 9; Length 599; Best Local Similarity 77.8%; Pred. No. 2.7e+02; Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	5 CTGCTTGTCAGACTTATCCAGATTCTTGGCGGTAG 40 437 CTGGATGTTGTAACCTATCTAGATTCTTTGGCGGTAG 472	599 bp DNA linear GSS 31-OCT-2002	599	599	599	599	599	599
ORIGIN										
FEATURES source	Location/Qualifiers 1..599 /organism="Candida glabrata" /mol_type="genomic DNA" /strain="CBS 138" /db_xref="taxon:5478" /clone="CG0802" /clone_lib="Candida glabrata Random Genomic Library"	Query Match 58.0%; Score 23.2; DB 9; Length 599; Best Local Similarity 77.8%; Pred. No. 2.7e+02; Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	5 CTGCTTGTCAGACTTATCCAGATTCTTGGCGGTAG 40 437 CTGGATGTTGTAACCTATCTAGATTCTTTGGCGGTAG 472	599 bp DNA linear GSS 31-OCT-2002	599	599	599	599	599	599
ORIGIN										
FEATURES source	Location/Qualifiers 1..599 /organism="Candida glabrata" /mol_type="genomic DNA" /strain="CBS 138" /db_xref="taxon:5478" /clone="CG0802" /clone_lib="Candida glabrata Random Genomic Library"	Query Match 58.0%; Score 23.2; DB 9; Length 599; Best Local Similarity 77.8%; Pred. No. 2.7e+02; Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	5 CTGCTTGTCAGACTTATCCAGATTCTTGGCGGTAG 40 437 CTGGATGTTGTAACCTATCTAGATTCTTTGGCGGTAG 472	599 bp DNA linear GSS 31-OCT-2002	599	599	599	599	599	599
ORIGIN										
FEATURES source	Location/Qualifiers 1..599 /organism="Candida glabrata" /mol_type="genomic DNA" /strain="CBS 138" /db_xref="taxon:5478" /clone="CG0802" /clone_lib="Candida glabrata Random Genomic Library"	Query Match 58.0%; Score 23.2; DB 9; Length 599; Best Local Similarity 77.8%; Pred. No. 2.7e+02; Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	5 CTGCTTGTCAGACTTATCCAGATTCTTGGCGGTAG 40 437 CTGGATGTTGTAACCTATCTAGATTCTTTGGCGGTAG 472	599 bp DNA linear GSS 31-OCT-2002	599	599	599	599	599	599
ORIGIN										
FEATURES source	Location/Qualifiers 1..599 /organism="Candida glabrata" /mol_type="genomic DNA" /strain="CBS 138" /db_xref="taxon:5478"<									

AUTHORS

Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L., Tsai, M., Cloutier, A., Lee, D., Gilm, N., Olson, T., Mayo, M., Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S., Schein, J., Marra, M., de Jong, P., Keele, J. W. and Kappes, S. M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)
Other_GSSs: CH240_406N8.T7

TITLE
JOURNAL
COMMENT

Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering-information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada.
Plate: 406 row: N column: 8
Seq primer: SP6
Class: BAC ends.

FEATURES

source

1. .716
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_406N8"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull LI Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

ORIGIN

Query Match 57.5%; Score 23; DB 9; Length 716;
Best Local Similarity 83.9%; Pred. No. 3.4e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGAGCTGCTTGTTCAGACTTATCCAGATTC 31

|||||
Db 517 AGAGCTGCTTGTTCAGACTTATCCAGACCT 487

RESULT 11

CC500299

LOCUS
DEFINITION
CH240_33811.TARBAC13P2 CHORI-240 Bos taurus genomic clone
CH240_33811, genomic survey sequence.

ACCESSION
CC500299.1 GI:31818551

VERSION

KEYWORDS

SOURCE

Bos taurus (cow)

ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

REFERENCE
1 (bases 1 to 807)

AUTHORS

Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L., Tsai, M., Cloutier, A., Lee, D., Gilm, N., Olson, T., Mayo, M., Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S., Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W., Dalrymple, B. P. and Tellam, R.
Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other_GSSs: CH240_33811.T7

TITLE

JOURNAL

COMMENT

Contact: Rob Holt
Sequencing

The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering-information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada.
Plate: 338 row: I column: 1
Seq primer: SP6
Class: BAC ends.

FEATURES

source

1. .807
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_33811"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull LI Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

ORIGIN

Query Match 57.5%; Score 23; DB 9; Length 807;
Best Local Similarity 83.9%; Pred. No. 3.5e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGAGCTGCTTGTTCAGACTTATCCAGATTC 31

|||||
Db 657 AGAGCTGCTTGTTCAGACTTATCCAGACCT 687

RESULT 12

AZ959690/c

LOCUS

DEFINITION
2M0227020F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0227020 F, genomic survey sequence.

ACCESSION
AZ959690

VERSION
AZ959690.1 GI:13830917

KEYWORDS

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0227 row: O column: 20

Seq primer: CGTGTAAACGACGGCCAGT

Class: plasmid ends

High quality sequence stop: 610.

Location/Qualifiers

FEATURES

source

1..610
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0227020"
 /sex="Female"
 /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC2M library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (GI|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 57.0%; Score 22.8; DB 9; Length 610;

Best Local Similarity 79.4%; Pred. No. 4e+02; Indels 0; Gaps 0;
 Matches 27; Conservative 0; Mismatches 7;

QY 1 AGAGCTGCTGGTCAGACTTATCCAGATTTCTGG 34

Db 474 AGAGTTCTTAGTCAGACTTCTACAGGTTCTTG 441

RESULT 13

AY105678/c

LOCUS

AY105678 PC0139313 mRNA linear HTC 22-FEB-2005

DEFINITION Zea mays PC0139313 mRNA sequence.

ACCESSION AY105678

VERSION AY105678.1 GI:21208756

KEYWORDS HTC.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 728)

Gardiner,J., Schroeder,S., Polacco,M.L., Sanchez-Villeda,H., Fang,Z., Morgante,M., Landewe,T., Fengler,K., Ueche,F., Hanafey,M., Tingey,S., Chou,H., Wing,R., Soderlund,C. and Coe,E.H. Jr.

Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization

Plant Physiol. 134 (4), 1317-1326 (2004)

15020742

2 (bases 1 to 728)

Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whittitt,M.S.,

Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

Maize Mapping Project/DuPont Consensus Sequences for Design of

Overgo Probes

Unpublished (2002)

3 (bases 1 to 728)

Coe,E.H.

TITLE JOURNAL

COMMENT

Direct Submission
 Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
 If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

source

Location/Qualifiers

1..728

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="maizeGDB:638121"

/db_xref="taxon:4577"

/clone_lib="Maize Mapping Project/DuPont Consensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Query Match 57.0%; Score 22.8; DB 4; Length 728;

Best Local Similarity 79.4%; Pred. No. 4.1e+02; Indels 0; Gaps 0;
 Matches 27; Conservative 0; Mismatches 7;

QY 4 GCTGCTTGCTGACACTTATCCAGATTTCTGGCGG 37

Db 305 GCAGCTTGCTGACACTTTCATGTTCTTGAGG 272

RESULT 14

DR799776/c

LOCUS

DR799776 ZM BF0022N18.r ZM_BFB Zea mays cDNA 5', mRNA linear EST 27-JUL-2005

DEFINITION DR799776

ACCESSION DR799776.1 GI:71326170

VERSION EST.

KEYWORDS Zea mays

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 771)

Kim,H., Collura,K., Wissotski,M., Smart,D., Kudrna,D., Muller,C.,

Rao,K., Haller,K., Wing,R., Soderlund,C., Walbot,V. and Yu,Y.

Maize Full-length cDNA Project

Unpublished (2005)

Contact: Yeisoo Yu

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ 85721-0036, USA

Tel: 520 626 9585

Fax: 520 621 1259

Email: yeisoo@genome.arizona.edu

Plate: 0022 row: N column: 18.

FEATURES

source

Location/Qualifiers

1..771

/organism="Zea mays"

/mol_type="mRNA"

/cultivar="B73"

/db_xref="taxon:4577"

/tissue type="mixed (silks, husks, ears, pollen, shoot tips, leaf, root tips, whole seed, embryo)"

/dev stage="varies by tissue"

/lab_host="DH10B T1 phage resistant"

/clone_lib="ZM_BFB"

/note="Vector: PCMV-SPORT 6.1; Site 1: EcoRV; Site 2: NotI; Maize Full length cDNA library (3530 library)"

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:28:37 ; Search time 73.5 Seconds
(without alignments)
967.380 Million cell updates/sec

Title: SEQ3-THEN-SEQ4

Perfect score: 40

Sequence: 1 agagtgctgtgcagcttaccgattcttggcgtag 40

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 0.1

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/1/ina/1_COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5_COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/H_COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq.*
- 7: /cgn2_6/prodata/1/ina/PP_COMB.seq.*
- 8: /cgn2_6/prodata/1/ina/RE_COMB.seq.*
- 9: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22.6	56.5	304533	3	US-09-949-016-15371 Sequence 15371, A
C 2	22.6	56.5	304533	3	US-09-949-016-15372 Sequence 15372, A
C 3	22.4	56.0	67479	3	US-09-949-016-11804 Sequence 11804, A
C 4	22.2	55.5	1600	3	US-09-938-540-1 Sequence 1, Appli
C 5	21.6	54.0	885	3	US-09-328-352-866 Sequence 866, App
C 6	21.6	54.0	254405	3	US-09-949-016-14381 Sequence 14381, A
C 7	21.2	53.0	282	2	US-08-440-725A-4 Sequence 4, Appli
C 8	21.2	53.0	282	6	PCT-US95-15991-4 Sequence 4, Appli
C 9	21.2	53.0	334	3	US-08-440-725A-7 Sequence 7, Appli
C 10	21.2	53.0	334	6	PCT-US95-15991-7 Sequence 7, Appli
C 11	20.8	52.0	601	3	US-09-949-016-167503 Sequence 167503, A
C 12	20.8	52.0	637	3	US-09-533-559-5575 Sequence 5575, Ap
C 13	20.8	52.0	58133	3	US-09-949-016-16464 Sequence 16464, A
C 14	20.8	52.0	96074	3	US-09-949-016-12760 Sequence 12760, A
C 15	20.8	52.0	96074	3	US-09-949-016-13611 Sequence 13611, A
C 16	20.8	52.0	110000	3	US-09-830-902-1 Sequence 1, Appli
C 17	20.6	51.5	2181	3	US-09-126-980-1 Sequence 1, Appli
C 18	20.6	51.5	2181	3	US-09-476-482-1 Sequence 1, Appli
C 19	20.6	51.5	2181	3	US-09-517-605-15 Sequence 15, Appl
C 20	20.4	51.0	683	3	US-09-533-559-6434 Sequence 6434, Ap
C 21	20.2	50.5	2337	3	US-09-934-899-7 Sequence 7, Appli
C 22	20.2	50.5	2337	3	US-09-934-868-27 Sequence 27, Appl
C 23	20.2	50.5	2337	3	US-10-701-200-27 Sequence 27, Appl
C 24	20	50.0	20	3	US-09-938-642-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-949-016-15371/c
; Sequence 15371, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15371
; LENGTH: 304533
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(304533)
; OTHER INFORMATION: n = A,T,C or G

Query Match 56.5%; Score 22.6; DB 3; Length 304533;
Best Local Similarity 75.7%; Pred. No. 19;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 AGAGCTGCTGTGCAGCTTACCAGATTCTTGGCGG 37
DB 255014 AGAGCTTCTGTTCGCTGAGCCAGATCTTGTGG 254978

RESULT 2

US-09-949-016-15372/c
; Sequence 15372, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307


```
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14381

Query Match      54.0%; Score 21.6; DB 3; Length 254405;
Best Local Similarity 75.0%; Pred. No. 50;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GAGCTGCTGGTCAGACTTATCCAGATTCTTGGCGG 37
Db 241977 GTGCTGCTGGTTAGAATTTGAAGATTCTTCTCAG 242012

RESULT 7
US-08-440-725A-4/c
; Sequence 4, Application US/08440725A
; Patent No. 5985822
; GENERAL INFORMATION:
; APPLICANT: Edelman, Gerald M.
; APPLICANT: Crossin, Kathryn L.
; APPLICANT: Sporns, Olaf
; APPLICANT: Krushel, Leslie
; TITLE OF INVENTION: INHIBITION OF GLIAL CELL PROLIFERATION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10550 No. 5985822th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,725A
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,658
; FILING DATE: 09-DEC-1994
; NAME: Holmes, Emily
; REGISTRATION NUMBER: 40,652
; REFERENCE/DOCKET NUMBER: TSRI 457.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 784-2937
; TELEFAX: (619) 784-9399
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-440-725A-4

Query Match      53.0%; Score 21.2; DB 2; Length 282;
Best Local Similarity 76.5%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CTGCTTGGTCAGACTTATCCAGATTCTTGGCGGT 38
Db 211 CTGCTTGGTCAGACTTATCCACCTTCTTGATGAT 178

RESULT 8
PCT-US95-15991-4/c
; Sequence 4, Application PC/TUS9515991
; Patent No. 5985822
; GENERAL INFORMATION:
; APPLICANT: Edelman, Gerald M.
; APPLICANT: Crossin, Kathryn L.
; APPLICANT: Sporns, Olaf
; APPLICANT: Krushel, Leslie
; TITLE OF INVENTION: INHIBITION OF GLIAL CELL PROLIFERATION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10550 No. 5985822th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15991
; FILING DATE: 11-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,658
; FILING DATE: 09-DEC-1994
; APPLICATION NUMBER: US 08/440,725
; FILING DATE: 08-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cepuritis, Talivaldis
; REGISTRATION NUMBER: 20,818
; REFERENCE/DOCKET NUMBER: TSRI 457.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-580-1180
; TELEFAX: 312-580-1189
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Gallus gallus
; TISSUE TYPE: Brain
PCT-US95-15991-4

Query Match      53.0%; Score 21.2; DB 6; Length 282;
Best Local Similarity 76.5%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CTGCTTGGTCAGACTTATCCAGATTCTTGGCGGT 38
Db 211 CTGCTTGGTCAGACTTATCCACCTTCTTGATGAT 178

RESULT 9
US-08-440-725A-7/c
; Sequence 7, Application US/08440725A
; Patent No. 5985822
; GENERAL INFORMATION:
; APPLICANT: Edelman, Gerald M.
; APPLICANT: Crossin, Kathryn L.
; APPLICANT: Sporns, Olaf
; APPLICANT: Krushel, Leslie
; TITLE OF INVENTION: INHIBITION OF GLIAL CELL PROLIFERATION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10550 No. 5985822th Torrey Pines Road, TPC-8
; CITY: La Jolla
```

```
/ STATE: California
/ COUNTRY: US
/ ZIP: 92037
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/440,725A
/ FILING DATE: 08-MAY-1995
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/353,658
/ FILING DATE: 09-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Holmes, Emily
/ REGISTRATION NUMBER: 40,652
/ REFERENCE/DOCKET NUMBER: TSRI 457.1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (619) 784-2937
/ TELEFAX: (619) 784-9399
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 334 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ US-08-440-725A-7

Query Match 53.0%; Score 21.2; DB 2; Length 334;
Best Local Similarity 76.5%; Pred. No. 12;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CTGCTTGGTCAGACTTATCCAGATTCTTGGCGGT 38
Db 225 CTGCTTGGTCAGACTTATCCAGATTCTTGGCGGT 192

RESULT 10
PCT-US95-15991-7/c
/ Sequence 7. Application PC/TUS9515991
/ GENERAL INFORMATION:
/ APPLICANT: Egeiman, Gerald M.
/ APPLICANT: Crossin, Kathryn L.
/ APPLICANT: Sporns, Olaf
/ APPLICANT: Krushel, Leslie
/ TITLE OF INVENTION: INHIBITION OF GLIAL CELL PROLIFERATION
/ NUMBER OF SEQUENCES: 7
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Olson & Hiersl, Ltd.
/ STREET: 20 North Wacker Drive Suite 3000
/ CITY: Chicago
/ STATE: IL
/ COUNTRY: USA
/ ZIP: 60606
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/15991
/ FILING DATE: 11-DEC-1995
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/353,658
/ FILING DATE: 09-DEC-1994
/ APPLICATION NUMBER: US 08/440,725
/ FILING DATE: 08-MAY-1995
```

```
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Cepuritis, Talivaldis
/ REGISTRATION NUMBER: 20,818
/ REFERENCE/DOCKET NUMBER: TSRI 457.1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312-580-1180
/ TELEFAX: 312-580-1189
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 334 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ PCT-US95-15991-7

Query Match 53.0%; Score 21.2; DB 6; Length 334;
Best Local Similarity 76.5%; Pred. No. 12;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CTGCTTGGTCAGACTTATCCAGATTCTTGGCGGT 38
Db 225 CTGCTTGGTCAGACTTATCCAGATTCTTGGCGGT 192

RESULT 11
US-09-949-016-167503/c
/ Sequence 167503, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL0001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 167503
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
/ US-09-949-016-167503

Query Match 52.0%; Score 20.8; DB 3; Length 601;
Best Local Similarity 70.0%; Pred. No. 21;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGAGCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
Db 203 AGATCAGCTGGGAGCTGACTTAAGCCTGAGTCTGGGCGGTGG 164

RESULT 12
US-09-533-559-5575/c
/ Sequence 5575, Application US/09533559
/ Patent No. 6902887
/ GENERAL INFORMATION:
/ APPLICANT: Randy M. Berka
/ APPLICANT: Michael W. Rey
/ APPLICANT: Jeffrey R. Shuster
/ APPLICANT: Sakari Kauppinen
/ APPLICANT: Ib Groth Clausen
/ APPLICANT: Peter Bjarke Olsen
/ TITLE OF INVENTION: Methods For Monitoring Multiple Gene
```


; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849,200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5575
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
US-09-533-559-5575

Query Match 52.0%; Score 20.8; DB 3; Length 637;
Best Local Similarity 70.0%; Pred. No. 21;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 AGAGCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
Db 308 AGGTTACTTGGTCACTGGTAAATAGAGTCTGGCGGTAG 269

RESULT 13

US-09-949-016-16464/C
; Sequence 16464, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16464
; LENGTH: 58133
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16464

Query Match 52.0%; Score 20.8; DB 3; Length 58133;
Best Local Similarity 70.0%; Pred. No. 75;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 AGAGCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
Db 12940 AGATCAGCTGGGACTGACTAGCCTGAGTCTGGCGCAGTGG 12901

RESULT 14

US-09-949-016-12760
; Sequence 12760, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12760
; LENGTH: 96074
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12760

Query Match 52.0%; Score 20.8; DB 3; Length 96074;
Best Local Similarity 78.1%; Pred. No. 86;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 9 TTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
Db 71820 TTGTGAGATTATCCAGATTGTTCAGAGTAG 71851

RESULT 15

US-09-949-016-13611
; Sequence 13611, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13611
; LENGTH: 96074
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13611

Query Match 52.0%; Score 20.8; DB 3; Length 96074;
Best Local Similarity 78.1%; Pred. No. 86;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 9 TTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
Db 71820 TTGTGAGATTATCCAGATTGTTCAGAGTAG 71851

Search completed: April 18, 2006, 19:31:16
Job time : 75.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:39:11 ; Search time 675.5 Seconds
(without alignments)
489.674 Million cell updates/sec

Title: SEQ3-THEN-SEQ4

Perfect score: 40

Sequence: 1 agagctgttggtcagacttaccagattcttggcggttag 40

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 0.1

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.Main.*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	24.4	61.0	1116	4	US-09-925-065A-14554 Sequence 14554, A
2	23.2	58.0	881	3	US-09-770-445-523 Sequence 523, App
3	22.6	56.5	413	8	US-10-723-860-495 Sequence 495, App
4	22.6	56.5	2091	7	US-10-451-467A-501 Sequence 501, App
C 5	22.2	55.5	1107	3	US-09-738-626-1385 Sequence 1385, App
C 6	22.2	55.5	1230	10	US-11-006-098-301 Sequence 301, App
C 7	22.2	55.5	1294	8	US-10-495-066-1 Sequence 1, Appli
C 8	22.2	55.5	1600	3	US-09-938-540-1 Sequence 1, Appli
C 9	22.2	55.5	1600	8	US-10-895-849-1 Sequence 1, Appli
C 10	22.2	55.5	3309400	3	US-09-738-626-1 Sequence 1, Appli
C 11	22	55.0	619	4	US-09-925-065A-14554 Sequence 14554, A
C 12	22	55.0	634	4	US-09-925-065A-269626 Sequence 269626, A
C 13	21.6	54.0	1544	4	US-09-925-065A-683526 Sequence 683526, A
C 14	21.6	54.0	400660	8	US-10-388-838-68 Sequence 68, Appl
C 15	21.4	53.5	408	7	US-10-767-701-19730 Sequence 19730, A
C 16	21.4	53.5	1296	7	US-10-182-243-16 Sequence 16, Appl
C 17	21.4	53.5	1628	9	US-10-425-114-26266 Sequence 26266, A
C 18	21.4	53.5	1842	9	US-10-450-763-23345 Sequence 23345, A
C 19	21.4	53.5	2378	5	US-10-764-420-40 Sequence 40, Appl
C 20	21.2	53.0	574	5	US-10-027-632-318758 Sequence 318758, A
C 21	21.2	53.0	574	6	US-10-027-632-318758 Sequence 318758, A
C 22	21.2	53.0	1412	8	US-10-425-115-166235 Sequence 166235, A
C 23	21	52.5	2291	4	US-09-925-065A-684988 Sequence 684988, A

24 52.5 2291 4 US-09-925-065A-684989 Sequence 684989, A
C 25 52.5 35143 5 US-10-087-192-1894 Sequence 1894, App
C 26 52.5 122186 3 US-09-563-728A-36 Sequence 36, Appl
27 20.8 247 7 US-10-424-599-20030 Sequence 20030, A
28 20.8 567 4 US-09-925-065A-347566 Sequence 347566, A
C 29 20.8 637 8 US-10-653-047-5575 Sequence 5575, App
C 30 20.8 912 7 US-10-282-122A-15945 Sequence 15945, A
C 31 20.6 369 7 US-10-282-122A-42393 Sequence 42393, A
C 32 20.6 422 4 US-09-925-065A-472479 Sequence 472479, A
C 33 20.6 51.5 422 4 US-10-151-274-15 Sequence 15, Appl
C 34 20.6 51.5 2181 5 US-09-932-257A-18 Sequence 18, Appl
C 35 20.6 51.5 2360 7 US-10-635-854A-1 Sequence 1, Appl
C 36 20.6 51.5 2360 9 US-10-756-149-2720 Sequence 2720, App
C 37 20.6 51.5 421609 7 US-10-367-094-122 Sequence 122, App
C 38 20.4 51.0 391 7 US-10-424-599-15268 Sequence 15268, A
C 39 20.4 51.0 444 7 US-10-021-323-16905 Sequence 16905, A
C 40 20.4 51.0 526 7 US-10-021-323-13421 Sequence 13421, A
C 41 20.4 51.0 543 3 US-09-563-817-591 Sequence 591, App
C 42 20.4 51.0 554 4 US-09-925-065A-426652 Sequence 426652, A
C 43 20.4 51.0 590 7 US-10-021-323-15374 Sequence 15374, A
C 44 20.4 51.0 600 8 US-10-425-115-63642 Sequence 63642, A
C 45 20.4 51.0 615 5 US-10-027-632-230216 Sequence 230216, A

ALIGNMENTS

RESULT 1
US-09-925-065A-14554
; Sequence 14554, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 557086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14554
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-14554

Query Match 61.0%; Score 24.4; DB 4; Length 1116;
Best Local Similarity 82.4%; Pred. No. 2.8;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 CTGCTTGTTCAGACTTATCCAGATTCTTGGCGGT 38
|||||
Db 582 CTGCTTGTTCAGATTATCCACATGCTTGCAGT 615
|||||

RESULT 2
US-09-770-445-523
; Sequence 523, Application US/09770445
; Patent No. US2002023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.

APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Matthew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Kriker, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
TITLE OF INVENTION: thaliana
FILE REFERENCE: 2023US (PARA-012PRV)
CURRENT APPLICATION NUMBER: US/09/770,445
PRIOR FILING DATE: 2001-01-26
CURRENT FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 523
LENGTH: 881
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(881)
OTHER INFORMATION: n = A,T,C or G
US-09-770-445-523

Query Match 58.0%; Score 23.2; DB 3; Length 881;
Best Local Similarity 77.8%; Pred. No. 9.2;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 AGCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGT 38
DB 177 AGCTGCAATGTCATCTACTATTCAGGTTCTTGGAGGT 212

RESULT 3
US-10-723-860-495
Sequence 495, Application US/10723860
Publication No. US20040253606A1
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Gineburg, Wendy M.
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &
TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
FILE REFERENCE: 05882.0193.NPUS01
CURRENT APPLICATION NUMBER: US/10/723,860
CURRENT FILING DATE: 2003-11-26
PRIOR FILING DATE: 2002-11-26
NUMBER OF SEQ ID NOS: 8393
SOFTWARE: PatentIn version 3.2
SEQ ID NO 495
LENGTH: 413
TYPE: DNA
ORGANISM: Homo sapiens
US-10-723-860-495

Query Match 56.5%; Score 22.6; DB 8; Length 413;
Best Local Similarity 75.7%; Pred. No. 14;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GAGCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGT 38
DB 177 AGCTGCAATGTCATCTACTATTCAGGTTCTTGGAGGT 212

Db 276 GAGCTGCTGTCGCCAGCTCATCCAGATTCTTGGTGGT 312

RESULT 4
US-10-451-467A-501
Sequence 501, Application US/10451467A
Publication No. US20040161840A1
GENERAL INFORMATION:
APPLICANT: CONTRERAS, ROLAND HENRI
APPLICANT: EBERHARDT, INES
APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
APPLICANT: REEKMAN, RIEKA JOSEPHINA
TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
TITLE OF INVENTION: YEAST AND FUNGI
FILE REFERENCE: JAB-1667
CURRENT APPLICATION NUMBER: US/10/451,467A
CURRENT FILING DATE: 2003-06-19
PRIOR APPLICATION NUMBER: EP 00870318.3
PRIOR FILING DATE: 2000-12-22
CURRENT FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: EP 01870002.1
PRIOR FILING DATE: 2001-01-04
CURRENT APPLICATION NUMBER: EP 01870003.9
PRIOR FILING DATE: 2001-01-09
NUMBER OF SEQ ID NOS: 732
SOFTWARE: PatentIn version 3.1
SEQ ID NO 501
LENGTH: 2091
TYPE: DNA
ORGANISM: Candida albicans
US-10-451-467A-501

Query Match 56.5%; Score 22.6; DB 7; Length 2091;
Best Local Similarity 75.7%; Pred. No. 21;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 GCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
DB 1870 GCTGTCAGACCATCATATCAAGATTATTGGAGGTTG 1906

RESULT 5
US-09-738-626-1385/c
Sequence 1385, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENO, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 1385
LENGTH: 1107
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-1385

Query Match 55.5%; Score 22.2; DB 3; Length 1107;
Best Local Similarity 88.9%; Pred. No. 27;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 CAGACTTATCCAGATTCTTGGCGGTAG 40
DB 704 CCGTCATATCCAGATTCTTGGCGGTAG 678

RESULT 6

US-11-006-098-301/c
; Sequence 301, Application US/11006098
; Publication No. US20050153402A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE OF INVENTION: REGULATORY PROTEINS
; FILE REFERENCE: BGI-123CP
; CURRENT APPLICATION NUMBER: US/11/006,098
; PRIOR FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US/09/602,874
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142690
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151251
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932206.6
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 362

Query Match 55.5%; Score 22.2; DB 8; Length 1294;
Best Local Similarity 88.9%; Pred. No. 28;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 CAGACTTATCCAGATTCTTGGCGGTAG 40
DB 861 CCGTCATATCCAGATTCTTGGCGGTAG 835

RESULT 8

US-09-938-540-1/c
; Sequence 1, Application US/09938540
; Patent No. US20020151001A1
; GENERAL INFORMATION:
; APPLICANT: Degussa AG
; TITLE OF INVENTION: New nucleotide sequences which code for the ccpA1 gene
; FILE REFERENCE: 000059 BT
; CURRENT APPLICATION NUMBER: US/09/938,540
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225)..(1388)
; OTHER INFORMATION: ccpA1-Gen
US-09-938-540-1

Query Match 55.5%; Score 22.2; DB 3; Length 1600;
Best Local Similarity 88.9%; Pred. No. 29;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 CAGACTTATCCAGATTCTTGGCGGTAG 40
DB 985 CCGTCATATCCAGATTCTTGGCGGTAG 959

RESULT 9

US-10-895-849-1/c
; Sequence 1, Application US/10895849
; Publication No. US20050032179A1
; GENERAL INFORMATION:
; APPLICANT: Degussa AG
; TITLE OF INVENTION: New nucleotide sequences which code for the ccpA1 gene
; FILE REFERENCE: 000059 BT
; CURRENT APPLICATION NUMBER: US/10/895,849
; CURRENT FILING DATE: 2004-07-22
; PRIOR APPLICATION NUMBER: US/09/938,540

RESULT 7

US-10-495-066-1/c
; Sequence 1, Application US/10495066
; Publication No. US20050014234A1
; GENERAL INFORMATION:
; APPLICANT: Zelder, Oskar
; APPLICANT: Pompejus, Markus
; APPLICANT: Schroder, Hartwig

Query Match 55.5%; Score 22.2; DB 10; Length 1230;
Best Local Similarity 88.9%; Pred. No. 27;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 CAGACTTATCCAGATTCTTGGCGGTAG 40
DB 804 CCGTCATATCCAGATTCTTGGCGGTAG 778

```

; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225)..(1388)
; OTHER INFORMATION: ccpA1-Gen
US-10-895-849-1

Query Match      55.5%; Score 22.2; DB 8; Length 1600;
Best Local Similarity 88.9%; Pred. No. 29;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      14 CAGACTTATCCAGATTCTTGGCGGTAG 40
Db      985 CCGTCATATCCAGATTCTTGGCGGTAG 959

RESULT 10
US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match      55.5%; Score 22.2; DB 3; Length 3309400;
Best Local Similarity 88.9%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      14 CAGACTTATCCAGATTCTTGGCGGTAG 40
Db      1317041 CCGTCATATCCAGATTCTTGGCGGTAG 1317015

RESULT 11
US-09-925-065A-381434/c
; Sequence 381434, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135

```

```

; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 381434
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-381434

Query Match      55.0%; Score 22; DB 4; Length 619;
Best Local Similarity 78.1%; Pred. No. 29;
Matches 25; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      9 TTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
Db      538 TTGGACAGACTAATACAGTAAGTTGGCRGTAG 507

RESULT 12
US-09-925-065A-269626
; Sequence 269626, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269626
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-269626

Query Match      55.0%; Score 22; DB 4; Length 634;
Best Local Similarity 73.7%; Pred. No. 29;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      2 GAGCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTA 39
Db      361 GAGCTTCTTTGTCAGTCTTACACAGAAATTTGTTGGAA 398

RESULT 13
US-09-925-065A-683526
; Sequence 683526, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

```

TITLE OF INVENTION: Identification and Mapping of Single
Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 683526
LENGTH: 1544
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-683526

Query Match 54.0%; Score 21.6; DB 4; Length 1544;
Best Local Similarity 75.0%; Pred. No. 54;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GAGCTGCTGGTCAGACTTATCCAGATTCTTGGCGG 37
DB 1124 GTGCTGCTGGTTAGATTGGAAGATTCTTCTCAG 1159

RESULT 14

US-10-388-838-68
Sequence 68, Application US/10388838
Publication No. US20040180344A1

GENERAL INFORMATION:
APPLICANT: David W. Morris

APPLICANT: Marc Malandro

TITLE OF INVENTION: Novel Therapeutic Targets in Cancer

FILE REFERENCE: 529452001600

CURRENT APPLICATION NUMBER: US/10/388,838

CURRENT FILING DATE: 2003-03-14

NUMBER OF SEQ ID NOS: 114

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 68

LENGTH: 400660

TYPE: DNA

ORGANISM: Homo sapiens

US-10-388-838-68

Query Match 54.0%; Score 21.6; DB 8; Length 400660;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GAGCTGCTGGTCAGACTTATCCAGATTCTTGGCGG 37
DB 379432 GTGCTGCTGGTTAGATTGGAAGATTCTTCTCAG 379467

RESULT 15

US-10-767-701-19730/c

Sequence 19730, Application US/10767701

Publication No. US20040172684A1

GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

Plants and Uses Thereof For Plant Improvement

FILE REFERENCE: 39-21(53535)B

CURRENT APPLICATION NUMBER: US/10/767,701

CURRENT FILING DATE: 2004-01-29

NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 19730
LENGTH: 408
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: LIB5049-002-R1-XPI-B3
US-10-767-701-19730

Query Match 53.5%; Score 21.4; DB 7; Length 408;
Best Local Similarity 71.8%; Pred. No. 48;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGAGCTGCTGGTCAGACTTATCCAGATTCTTGGCGGTA 39
DB 234 ATAGCTGACTGGGCGAGTCGCATCAAGATCTTCCGCGGA 196

Search completed: April 18, 2006, 23:14:02
Job time : 680.5 secs

This Page Blank (uspto)

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

- Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
- Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:49:40 ; Search time 352 Seconds
(without alignments)
457.708 Million cell updates/sec

Title: SEQ3-THEN-SEQ4

Perfect score: 40

Sequence: 1 agagctgctgttcagacttatccagattcttggcggtag 40

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.1

Searched: 9281099 seqs, 2013915447 residues

Total number of hits satisfying chosen parameters: 18562198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.New.*

1: /SIDSS/ptodata/2/pubpna/US06_NEW_PUB.seq.*
2: /SIDSS/ptodata/2/pubpna/US08_NEW_PUB.seq.*
3: /SIDSS/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /SIDSS/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /SIDSS/ptodata/2/pubpna/US05_NEW_PUB.seq.*
6: /SIDSS/ptodata/2/pubpna/US09_NEW_PUB.seq1.*
7: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq.*
8: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq1.*
9: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq2.*
10: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq3.*
11: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq1.*
12: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
13: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
14: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq4.*
15: /SIDSS/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24.4	61.0	1116	6	US-09-925-065A-14554
2	24.4	61.0	1116	9	US-10-301-480-115791
3	24.4	61.0	1116	10	US-10-301-480-729200
4	24.2	60.5	145985	9	US-10-330-773-140
5	22	55.0	589	9	US-10-301-480-72442
6	22	55.0	589	10	US-10-301-480-685851
7	22	55.0	619	6	US-09-925-065A-381434
8	22	55.0	621	10	US-10-301-480-450514
9	22	55.0	621	10	US-10-301-480-1063923
10	22	55.0	634	6	US-09-925-065A-269626
11	21.6	54.0	1544	6	US-09-925-065A-683526
12	21	52.5	2291	6	US-09-925-065A-684988
13	21	52.5	2291	6	US-09-925-065A-684989
14	20.8	52.0	567	6	US-09-925-065A-347566
15	20.8	52.0	611	10	US-10-301-480-419766
16	20.8	52.0	611	10	US-10-301-480-1033175
17	20.8	52.0	110000	14	US-11-155-492-1
18	20.6	51.5	422	6	US-09-925-065A-472479

19	20.6	51.5	600	8	US-10-750-185-2619	Sequence 2619, Ap
20	20.6	51.5	600	8	US-10-750-623-2619	Sequence 2619, Ap
21	20.4	51.0	554	6	US-09-925-065A-426652	Sequence 426652,
22	20.4	51.0	559	10	US-10-301-480-490105	Sequence 490105,
23	20.4	51.0	559	10	US-10-301-480-1103514	Sequence 1103514,
24	20.4	51.0	659	6	US-09-925-065A-872783	Sequence 872783,
25	20.4	51.0	988	10	US-10-301-480-547116	Sequence 547116,
26	20.4	51.0	988	10	US-10-301-480-1160525	Sequence 1160525,
27	20.4	51.0	1678	5	US-09-978-360A-186	Sequence 186, App
28	20.4	51.0	2431	8	US-10-750-185-42399	Sequence 42399, A
29	20.4	51.0	2431	8	US-10-750-623-42399	Sequence 42399, A
30	20.2	50.5	606	6	US-09-925-065A-868797	Sequence 868797,
31	20	50.0	596	6	US-09-925-065A-742008	Sequence 742008,
32	20	50.0	606	6	US-09-925-065A-774063	Sequence 774063,
33	20	50.0	630	9	US-10-932-182A-4131	Sequence 4131, Ap
34	20	50.0	630	9	US-10-932-182A-4131	Sequence 4131, Ap
35	20	50.0	732	10	US-10-301-480-547094	Sequence 547094,
36	20	50.0	732	10	US-10-301-480-1160503	Sequence 1160503,
37	20	50.0	999	10	US-10-301-480-560784	Sequence 560784,
38	20	50.0	999	10	US-10-301-480-1174193	Sequence 1174193,
39	20	50.0	1372	6	US-09-925-065A-708366	Sequence 708366,
40	19.8	49.5	527	6	US-09-925-065A-36314	Sequence 36314, A
41	19.8	49.5	527	6	US-09-925-065A-36315	Sequence 36315, A
42	19.8	49.5	527	9	US-10-301-480-137552	Sequence 137552,
43	19.8	49.5	527	9	US-10-301-480-137553	Sequence 137553,
44	19.8	49.5	527	10	US-10-301-480-750961	Sequence 750961,
45	19.8	49.5	527	10	US-10-301-480-750962	Sequence 750962,

ALIGNMENTS

RESULT 1
US-09-925-065A-14554
; Sequence 14554, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14554
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-14554

Query Match 61.0%; Score 24.4; DB 6; Length 1116;
Best Local Similarity 82.4%; Pred. No. 0.71;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5 CTCCTGGTCCAGATTATCCAGATTCTTGGCGGT 38
Db 582 CTCCTGGTCCAGATTATCCAGATTCTTGGCGGT 615

RESULT 2

US-10-301-480-115791
; Sequence 115791, Application US/10301480


```
QY 2 GAGCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTA 39
    ||||| ||| ||||| ||||| ||||| ||||| |||||
Db 361 GAGCTTCTTTGTGAGCTTATACAGAAATTTGTTGGA 398
    ||||| ||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-09-925-065A-683526
; Sequence 683526, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 683526
; LENGTH: 1544
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-683526

Query Match 54.0%; Score 21.6; DB 6; Length 1544;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GAGCTGCTTGGTCAGACTTATCCAGATTCTTGGCGG 37
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1124 GTGCTGCTTGGTTAGAAATTTGAAGATTCTTCTCAG 1159
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-09-925-065A-684988
; Sequence 684988, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 684988
; LENGTH: 2291
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-684988

Query Match 52.5%; Score 21; DB 6; Length 2291;
Best Local Similarity 73.0%; Pred. No. 33;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 4 GCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 491 GCTTCATGCTCAGAATTACTCATTCTTGTGGCAG 527
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-09-925-065A-684989
; Sequence 684989, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 684989
; LENGTH: 2291
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-684989

Query Match 52.5%; Score 21; DB 6; Length 2291;
Best Local Similarity 73.0%; Pred. No. 33;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 GCTGCTTGGTCAGACTTATCCAGATTCTTGGCGGTAG 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 491 GCTTCATGCTCAGAATTACTCATTCTTGTGGCAG 527
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-925-065A-347566
; Sequence 347566, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347566
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-347566

Query Match 52.5%; Score 21; DB 6; Length 2291;
Best Local Similarity 73.0%; Pred. No. 33;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

US-09-925-065A-347566

Query Match 52.0%; Score 20.8; DB 6; Length 567;
Best Local Similarity 78.1%; Pred. No. 27;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 8 CTTGGTCAGACTTATCCAGATTCTTGGCGGTA 39
|||||
Db 370 CTTGCTCAGAATCATCCAGATTTTAGGCCATA 401

RESULT 15

US-10-301-480-419766
; Sequence 419766, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 10827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 419766
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-419766

Query Match 52.0%; Score 20.8; DB 10; Length 611;
Best Local Similarity 78.1%; Pred. No. 28;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 8 CTTGGTCAGACTTATCCAGATTCTTGGCGGTA 39
|||||
Db r 371 CTTGCTCAGAATCATCCAGATTTTAGGCCATA 402

Search completed: April 18, 2006, 20:01:34
Job time : 353 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

QM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:11:02 ; Search time 948 Seconds
(without alignments)
2398.458 Million cell updates/sec

Title: SEQ4-THEN-SEQ3
Perfect score: 40

Sequence: 1 attcagattcttcggtagagagctgtgtcagactt 40

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.1

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sta.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	24.6	61.5	219245	14 AC130576	AC130576 Rattus no
C 2	24.6	61.5	222077	14 AC114229	AC114229 Rattus no
C 3	24.6	61.5	234153	14 AC114507	AC114507 Rattus no
C 4	24.4	61.0	110000	15 CR382128_09	Continuation (10 o
C 5	24	60.0	2902	1 SO23SRN	X68427 S.oralis ge
C 6	23.8	59.5	871	6 BD147349	BD147349 Primer fo
C 7	23.8	59.5	871	6 AX867287	AX867287 Sequence
C 8	23.8	59.5	2126	8 HSA505016	AJ505016 Homo sapi
C 9	23.8	59.5	2343	6 BD158115	BD158115 Primer fo
C 10	23.8	59.5	2343	6 AX879955	AX879955 Sequence
C 11	23.8	59.5	2343	8 AK022924	AK022924 Homo sapi
C 12	23.8	59.5	2390	6 BD156054	BD156054 Primer fo
C 13	23.8	59.5	2390	6 AX876294	AX876294 Sequence
C 14	23.8	59.5	2390	8 AK001233	AK001233 Homo sapi
C 15	23.8	59.5	2650	6 BD249583	BD249583 DNA repli
C 16	23.8	59.5	2763	8 BC096758	BC096758 Homo sapi
C 17	23.8	59.5	2823	8 BC018694	BC018694 Homo sapi
C 18	23.8	59.5	2892	6 BD159594	BD159594 Primer fo

C 19	23.8	59.5	2892	6 AX882205	AX882205 Sequence
C 20	23.8	59.5	2892	8 AK022573	AK022573 Homo sapi
C 21	23.8	59.5	3409	8 AK025149	AK025149 Homo sapi
C 22	23.8	59.5	3475	8 AK123857	AK123857 Homo sapi
C 23	23.8	59.5	6701	6 BD158653	BD158653 Primer fo
C 24	23.8	59.5	6701	6 AX880894	AX880894 Sequence
C 25	23.8	59.5	6701	8 AK023826	AK023826 Homo sapi
C 26	23.8	59.5	7150	8 AB007931	AB007931 Homo sapi
C 27	23.8	59.5	7228	6 AX677323	AX677323 Sequence
C 28	23.8	59.5	7827	6 AR338613	AR338613 Sequence
C 29	23.8	59.5	8825	6 AX780115	AX780115 Sequence
C 30	23.8	59.5	10852	6 CQ729868	CQ729868 Sequence
C 31	23.8	59.5	15552	6 CQ873939	CQ873939 Sequence
C 32	23.8	59.5	15552	8 AF348492	AF348492 Homo sapi
C 33	23.6	59.0	114934	5 BX890577	BX890577 Zebrafish
C 34	23	57.5	5190	9 AB103331	AB103331 Mus muscu
C 35	23	57.5	171231	14 AC127867	AC127867 Rattus no
C 36	23	57.5	178391	9 AC154141	AC154141 Mus muscu
C 37	23	57.5	205427	9 AC154658	AC154658 Mus muscu
C 38	23	57.5	218468	14 AC154737	AC154737 Mus muscu
C 39	23	57.5	245080	14 AC108967	AC108967 Rattus no
C 40	22.8	57.0	1585	15 D89263	D89263 Schizosacch
C 41	22.8	57.0	13936	15 SPAC1556	AL132984 S.pombe c
C 42	22.8	57.0	189425	14 AC142182	AC142182 Rattus no
C 43	22.8	57.0	223728	14 AC135443	AC135443 Rattus no
C 44	22.8	57.0	231785	14 AC094057	AC094057 Rattus no
C 45	22.6	56.5	1103	15 AK071737	AK071737 Oryza sat

ALIGNMENTS

RESULT 1

AC130576/c

LOCUS

DEFINITION

AC130576

AC130576.4

HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

Rattus norvegicus (Norway rat)

Rattus norvegicus

ORGANISM

Eukaryota; Metazoa;

Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Rattus;

1 (bases 1 to 219245)

REFERENCE

AUTHORS

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

AC130576 219245 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-283D7, WORKING DRAFT SEQUENCE.

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokeneme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Flopper, F., Polindexter, A., Popovic, D., Primus, E., Fu, L.-L., Puafo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valasek, R., Vera, V., Villaesca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission

2 (bases 1 to 219245)

Worley, K.C.

Direct Submission

Submitted (12-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 219245)

Direct Submission

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:23912386. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBBJ
Center clone name: CH230-283D7
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 202069 bases at least Q40
Consensus quality: 204430 bases at least Q30
Consensus quality: 205755 bases at least Q20
Estimated insert size: 207591; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced

* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 219245: contig of 219245 bp in length.

FEATURES

source

1..219245

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-283D7"

1..1173

/note="wgs_end_extension"

clone_end:Sp6"

1923..2795

/note="clone boundary"

clone_end:Sp6"

site:

end_sequence:BZ257707"

211604..212304

/note="clone boundary"

clone_end:T7

site:

end_sequence:BZ257705"

217179..219245

/note="wgs_end_extension"

clone_end:T7"

ORIGIN

Query Match 61.5%; Score 24.6; DB 14; Length 219245;
Best Local Similarity 76.9%; Pred. No.53;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATCCGATTCTTGGCGGTAGAGAGTGTGTCGACACT 39
|||||
DB 172592 ATCCGATTCTTGGCGGTAGACACAGTGGCCCACT 172554
|||||

RESULT 2

AC114229

LOCUS

DEFINITION

AC114229

AC114229.3

GI:2195356

HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 222077)

Muzny, D.Marie., Metzker, M.Lee., Abruzzo, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Johnson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwa, L., Loulseg, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, G., Muniz, S., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Pal, S., Parks, K., Pastermak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartabeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Shedd, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 222077)
Worley, K. C.

Direct Submission
Submitted (07-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 222077)

Direct Submission
Submitted (10-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Sep 19, 2002 this sequence version replaced gi:21744222.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: QWR
Center clone name: CH230-322D21
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 199431 bases at least Q40
Consensus quality: 201398 bases at least Q30
Consensus quality: 202516 bases at least Q20
Estimated insert size: 214381; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 10381: contig of 10381 bp in length
* 10382 10481: gap of unknown length
* 10482 220209: contig of 209728 bp in length
* 220210 220309: gap of unknown length
* 220310 222077: contig of 1768 bp in length.

FEATURES
source
1..222077
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-322D21"
3302..3950
/note="clone boundary
clone_end:Sp6
site:Mbol
end_sequence:RXANO23TV"
9359..10381
/note="wgs contig"
9946..64122
/note="clone boundary
clone_end:T7
site:Mbol
end_sequence:RXANO23TV"
10382..10481
/estimated_length=unknown
220210..220309
/estimated_length=unknown

misc_feature
misc_feature
misc_feature

gap
gap

ORIGIN
Query Match 61.5%; Score 24.6; DB 14; Length 222077;
Best Local Similarity 76.9%; Pred. No. 53;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATCCGATTTCTTGGCGGTAGAGAGCTGTGTGTCAGACT 39
|||||
DB 25944 ATCCGATTTCTTGGCGGTAGACACAGCTTGCCACACT 25982
|||||

RESULT 3
AC114507/c
LOCUS
DEFINITION
AC114507
Rattus norvegicus clone CH230-30113, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC114507
HTG: HTGS_PHASE1; HTGS DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.
1 (bases 1 to 234153)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alshrocks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falla, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

COMMENT

Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlesbird, D., Jackson, A., Jackson, L., Jacob, L. J., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowig, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhera, L., Loulsegod, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nair, L., Nwankweli, O., Okwuonu, G., Olarnpunasagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivas, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, K., White, F., Williams, G., Willison, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, J., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

2 (bases 1 to 234153)

Unpublished

Worley, K.C.

REFERENCE

JOURNAL

Submitted (10-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 234153)

Unpublished

Worley, K.C.

REFERENCE

JOURNAL

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:22857240. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNCO
Center clone name: CH230-30113
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 216821 bases at least Q40
Consensus quality: 219385 bases at least Q30

Consensus quality: 220753 bases at least Q20
Estimated insert size: 22569; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 228678: contig of 228678 bp in length
* 228679 228778: gap of unknown length
* 228779 231809: contig of 3031 bp in length
* 231810 231909: gap of unknown length
* 231910 234153: contig of 2244 bp in length.

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

* 1 234153
* /organism="Rattus norvegicus"
* /mol_type="Genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-30113"
* .1445
* /note="wgs contig"
* 1496. 3182
* /note="wgs contig"
* 227041. 228678
* /note="wgs contig"
* 228679. 228778
* /estimated_length=unknown
* 231810. 231909
* /estimated_length=unknown

```
CR382128-22 2200001 2310000
CR382128-23 2300001 2410000
CR382128-24 2400001 2510000
CR382128-25 2500001 2610000
CR382128-26 2600001 2710000
CR382128-27 2700001 2810000
CR382128-28 2800001 2910000
CR382128-29 2900001 3010000
CR382128-30 3000001 3066374
Continuation (10 of 31) of CR382128 Yarrowia lipolytica chrom

Query Match 61.0%; Score 24.4; DB 15; Length 110000;
Best Local Similarity 82.4%; Pred. No. 65;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TCACAGATCTTGGCGGTAGAGACTGCTTGGTCA 35
|||||
Db 17273 TCACAGATCTTGGAGTGGTGAGACTTGGTCA 17306

RESULT 5
SO23SRRN 5023SRRN 2902 bp DNA linear BCT 29-MAR-1993
LOCUS S.oralis gene for 23S rRNA.
DEFINITION X68427
ACCESSION X68427.1 GI:288522
VERSION 23S ribosomal RNA.
KEYWORDS Streptococcus oralis
SOURCE Streptococcus oralis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2902)
AUTHORS Ludwig, W., Kirchof, G., Klugbauer, N., Weizenegger, M., Betzl, D.,
Ehrmann, M., Hertel, C., Jilg, S., Tatzel, R., Zitzelsberger, H.,
Liebl, S., Hochberger, M., Shah, J., Lane, D. and Wallnoef, P.R.
TITLE Complete 23S ribosomal RNA sequences of Gram-positive Bacteria with
a low DNA G+C content
JOURNAL Syst. Appl. Microbiol. 15, 487-501 (1992)
REFERENCE 2 (bases 1 to 2902)
AUTHORS Ludwig, W.
TITLE Direct Submission
JOURNAL Submitted (08-SEP-1992) W. Ludwig, Lehrst. fuer Mikrobiologie TU
Muenchen, Arcistr. 21, 8000 Muenchen 2, FRG
FEATURES
source
1. .2902
/organism="Streptococcus oralis"
/mol_type="genomic DNA"
/strain="DSM 20066"
/db_xref="taxon:1303"
rRNA 1. .2902
/product="23S ribosomal RNA"
ORIGIN
Query Match 60.0%; Score 24; DB 1; Length 2902;
Best Local Similarity 84.4%; Pred. No. 1e+02;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATCCAGATCTTGGCGGTAGAGACTGCTTGG 32
|||||
Db 876 ATCAAGATCTTGGAGTGGTAGCAGCTTTGG 907

RESULT 6
BD147349/c BD147349 871 bp DNA linear PAT 17-JAN-2003
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD147349
ACCESSION BD147349.1 GI:27853107
VERSION JP 2002191363-A/2192.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
Hominidae; Homo.
REFERENCE 1 (bases 1 to 871)
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 2192 09-JUL-2002;
HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/2192
PD 09-JUL-2002
PF 28-JUL-2000 JP 2002080990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof PH Key
FT source 1. .871
FT Location/Qualifiers
/organism="Homo sapiens (human)"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
1. .871
Location/Qualifiers
Query Match 59.5%; Score 23.8; DB 6; Length 871;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATCTTGGCGGTAGAGACTGCTTGGTCA 35
|||||
Db 649 ATCCAATCTTGGCGGTAGGATGCTTTTCA 615

RESULT 7
AX867287/c AX867287 871 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 2192 from Patent EP1074617.
DEFINITION AX867287
ACCESSION AX867287.1 GI:40021669
VERSION
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 2192 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
1. .871
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 59.5%; Score 23.8; DB 6; Length 871;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATCTTGGCGGTAGAGACTGCTTGGTCA 35
|||||
Db 649 ATCCAATCTTGGCGGTAGGATGCTTTTCA 615
```

RESULT 8	BD158115/c	BD158115	2343 bp	DNA	linear	PAT 17-JAN-2003
LOCUS	Primer for synthesizing full-length cDNA and use thereof.					
DEFINITION	HSA05016	BD158115				
ACCESSION	Homo sapiens mRNA for retinoblastoma-associated factor 600-like					
VERSION	protein (RBAF600 gene), mutant allele.					
KEYWORDS	AJ505016	BD158115.1	GI:27863873			
SOURCE	RBAF600 gene; retinoblastoma-associated factor 600-like protein.					
ORGANISM	Homo sapiens (human)					
REFERENCE	1					
AUTHORS	Gentilini, C., Fatho, M., Lennerz, V., Lifke, A., Woelfel, C. and Woelfel, T.					
TITLE	Response of autologous T cells to a human melanoma is dominated by individual mutant antigens					
JOURNAL	Unpublished					
AUTHORS	Lennerz, V.					
TITLE	Direct Submission					
JOURNAL	Submitted (06-AUG-2002) Lennerz V., III. Medizinische Klinik, Johannes Gutenberg-University of Mainz, Langenbeckstrasse 1, 55101 Mainz, GERMANY					
COMMENT	Related sequence: AF348492.					
FEATURES	Location/Qualifiers					
source	1..2126					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/sex="female"					
	/cell_line="MZ7-MEL"					
	/cell_type="melanoma"					
	/country="Germany"					
gene	1..2126					
	/gene="RBAF600"					
CDS	41..1807					
	/gene="RBAF600"					
	/notes="mutant"					
	/codon_start=1					
	/evidence=experimental					
	/product="retinoblastoma-associated factor 600-like protein"					
	/protein_id="CAD43719.1"					
	/db_xref="GI:24474789"					
	/translation="MLDQINSTFVRSNPVLQGLRIIPYLSFGVEKMQILVERFK PYCNFDKVDHSGDDKVPFLDCFKIAAGIKNNSGHQIKDILQKIGITONALDYMKK HPSAKNLDADTWKFLSPALPFFILRLARGIAIHPGQVGLIGTDSIPNLHKLQVVS SDEGIGTAEINLEALREHPDYNKKIDARRETRAERKEMAMARQKALGTILGWTNE KQGVTKTALLKMELEIEPGLTCCI CREGYKFPQTQVIGIYTFTRVALBELNKP RQQGYSTVSHNIVHYDCHLAARLARGREWSAALQNKNGLLPVMRPHVE SAFATLARHNTYLOECTQREPTQLNTHDKLFLRFAMEQSFSDTGTGGGGRSNI HLPIYIHTVLVLTNRATSRKLNQFLQPKKVKVESAFVDGPPYFTVLALHI LPQQRATRVILRLIATVTSARAVACGATRLTDKAVQDYSAVRSILFWALVDLI YNMFKVPVSTNEGWSCLSIYIRHNDMPYEADKALKTFQREMPVETSEFLDV AGLLSEITDPESFLKDLLNSVP"					
variation	1025					
	/gene="RBAF600"					
	/replace="g"					
ORIGIN						
Query Match	59.5%;	Score 23.8;	DB 8;	Length 2126;		
Best Local Similarity	80.0%;	Pred. No. 1.3e+02;				
Matches	28;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;	
QY	1	ATCCAGATTCTTGGCGGTAGAGCTGCTTGGTCA	35			
Db	373	ATCCAAATCTTGGCGCTAGGATGTGCTTTTCA	339			
RESULT 9						
LOCUS	AX879955/c	AX879955	2343 bp	DNA	linear	PAT 17-DEC-2003
DEFINITION	Sequence 14860 from Patent EP1074617.					
ACCESSION	AX879955					
VERSION	AX879955.1	GI:40034691				
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	1					
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
TITLE	Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.					
JOURNAL	Primers for synthesizing full-length cDNA and their use					
FEATURES	Location/Qualifiers					
source	1..2343					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
	61..2025					
	/note="unnamed protein product"					

PC	C12P21/02.C12Q1/68//C12P21/08.G06F17/30.C12N15/00.C12N5/00	CC
Primer for synthesizing full-length cDNA and use thereof FH Key		
FT	CDS	Location/Qualifiers
(119)...(2080).		
FEATURES	Location/Qualifiers	
source	1..2390	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
ORIGIN		
Query Match	59.5%; Score 23.8; DB 6; Length 2390;	
Best Local Similarity	80.0%; Pred. No. 1.3e+02;	
Matches	28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
QY	1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGTC 35	
Db	649 ATCCAAATCTTGGCGGTAGAGATGCTTTTCA 615	
RESULT 13		
AX876294/c		
LOCUS	AX876294	2390 bp DNA linear PAT 17-DEC-2003
DEFINITION	Sequence 11199 from Patent EP1074617.	
ACCESSION	AX876294	
VERSION	AX876294.1	GI:40031030
KEYWORDS		
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
	Hominidae; Homo.	
REFERENCE	1	
AUTHORS	Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,	
	Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.	
TITLE	Primers for synthesizing full-length cDNA and their use	
JOURNAL	Patent: EP 1074617-A 11199 07-FEB-2001;	
	Research Association for Biotechnology (JP)	
FEATURES	Location/Qualifiers	
source	1..2390	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
CDS	119_2083	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAF89207.1"	
	/db_xref="GI:40031031"	
	/translation="MNTLVMLGTLNLVAEQSKDSGGAVAEQVLSIMBILDES	
	NAEPLSEKGNLLTGDKDQVLMDDINGTFVRSNPVSLQGLRIIPVLSFGEVEKM	
	QILVERPKYCNFDKDEHSGDDKVFDFCFKIAAGIKNSNGHQLKDLILQKITQ	
	NALDYNKKHPSAKNLADDIWKKFUSRPALPFIILRLGLAIOHPTQVLITGDSIPN	
	LHLKEQVSDGEGITAEALLREHPDNKKIDAARETRAEKKRWAMMRQKALG	
	TLCMTNTEGVVTKTALLQWBELEIEEPGLTCCICREGYKFPKTVLGIYFTKRV	
	LEELNKPQKQYSTVSHFNIVHYDCHLAAVRLARGREESALONANTKCNGLLP	
	VNGHPVPESAPATCLAHNTYLQECTGQREPTVQLNIHDIKLLEFPAMEQSFADTG	
	GGGRENHILPIYIITVLYLVNTRATREKNLQGFLEQPKGWESAFAVDGPYY	
	FTVALHILPPQWRAVRTEILRLVTSQARAVPGGATRLDKAVKDYSAYRSLL	
	TFSEFLDVAGLLSEITDPESFLKDLNSVP"	
ORIGIN		
Query Match	59.5%; Score 23.8; DB 6; Length 2390;	
Best Local Similarity	80.0%; Pred. No. 1.3e+02;	
Matches	28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
QY	1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGTC 35	
Db	649 ATCCAAATCTTGGCGGTAGAGATGCTTTTCA 615	
RESULT 14		
AK001233/c		

LOCUS	AK001233	2390 bp mRNA linear PRI 30-JAN-2004
DEFINITION	Homo sapiens cDNA FLJ10371 fis, clone NT2RM2001903, highly similar to Homo sapiens mRNA for KIAA0462 protein.	
ACCESSION	AK001233	
VERSION	AK001233.1	GI:7022359
KEYWORDS	oligo capping; fis (full insert sequence).	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
	Hominidae; Homo.	
REFERENCE	1	
AUTHORS	Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,	
	Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,	
	Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,	
	Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,	
	Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,	
	Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,	
	Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furiya, T.,	
	Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,	
	Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,	
	Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,	
	Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,	
	Yosida, M., Houta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,	
	Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,	
	Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,	
	Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,	
	Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,	
	Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,	
	Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,	
	Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,	
	Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,	
	Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,	
	Ozaki, K., Hiroo, M., Ohmori, Y., Kawabata, A., Hiki, J., Kobatake, N.,	
	Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,	
	Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K.,	
	Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,	
	Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,	
	Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,	
	Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, K.,	
	Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.	
	Complete sequencing and characterization of 21,243 full-length	
	human cDNAs	
	Nat. Genet. 36 (1), 40-45 (2004)	
JOURNAL	14702039	
PUBMED		
REFERENCE	2	
AUTHORS	Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,	
	Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,	
	Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,	
	Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,	
	Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,	
	Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,	
	Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.	
	NEDO human cDNA sequencing project	
TITLE	Unpublished	
JOURNAL	3 (bases 1 to 2390)	
REFERENCE	Isogai, T. and Otsuki, T.	
AUTHORS	Direct Submission	
TITLE	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,	
JOURNAL	Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
	(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)	
	NEDO human cDNA sequencing project supported by Ministry of	
	International Trade and Industry of Japan; cDNA full insert	
	sequencing: Research Association for Biotechnology; cDNA library	
	construction, 5'- & 3'-end one pass sequencing and clone selection:	
	Helix Research Institute (supported by Japan Key Technology Center	
	etc.) and Department of Virology, Institute of Medical Science,	
	University of Tokyo.	
FEATURES	Location/Qualifiers	
source	1..2390	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	

ORIGIN

Query Match 59.5%; Score 23.8; DB 8; Length 2390;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35
|||||
Db 883 ATCCAAATTCTTGGCGCTAGGATGTCITTTTCA 849
/clone="NT2RM2001903"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM2"
/notes="cloning vector: pME18SFL3
mRNA from uninduced NT2 neuronal precursor cells."

Search completed: April 18, 2006, 19:42:50
Job time : 953 secs

RESULT 15

BD249583/c

LOCUS 2650 bp DNA linear PAT 17-JUL-2003

DEFINITION DNA replication- and repair-associated proteins.

ACCESSION BD249583

VERSION BD249583.1 GI:33059353

KEYWORDS JP 2002525032-A/3.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 2650)
Tang,T.Y., Corley,N.C., Baughn,M.R., Reddy,R., Guegler,K.J. and
Yue,H.
DNA replication- and repair-associated proteins
Patent: JP 2002525032-A 3 13-AUG-2002;
INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002525032-A/3
PD 13-AUG-2002
PF 06-AUG-1999 JP 2000563781
PR 07-AUG-1998 US 60/155245,27-AUG-1998 US 60/155181 PI
TOM Y TANG,NEIL C CORLEY,MARIAH R BAUGHN,ROOPA REDDY,KARL J PI
GUEGLER,
PI HENRY YUE
PC C12N15/09,A61K38/00,A61K45/00,A61P1/04,A61P1/16,A61P5/06 PC
,A61P5/14,A61P7/00,
PC A61P7/06,A61P9/10,A61P11/00,A61P11/06,A61P13/00,A61P13/12, PC
A61P15/00,
PC A61P17/00,A61P19/00,A61P19/08,A61P21/00,A61P21/04,A61P25/00,
PC A61P27/06,
PC A61P27/12,A61P29/00,A61P29/00,A61P31/04,A61P31/10,A61P31/12,
PC A61P31/18,
PC A61P33/00,A61P35/00,A61P35/02,A61P37/00,A61P37/08,A61P43/00,
PC C07K14/47,
PC C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/02 PC
,C12Q1/68,G01N33/53,
PC G01N33/566//G01N33/15,G01N33/50,(C12P21/02,C12R1:91),(C12P21/
PC 02,C12R1:19),
PC C12N15/00,C12N5/00,A61K37/02
CC Incyte Clone No: 1849911
FH Key Location/Qualifiers
FT source 1..2650
/organism="Homo sapiens (human)".
FT Location/Qualifiers
1..2650
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source

ORIGIN

Query Match 59.5%; Score 23.8; DB 6; Length 2650;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35
|||||
Db 883 ATCCAAATTCTTGGCGCTAGGATGTCITTTTCA 849

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:05:32 ; Search time 219 Seconds
(without alignments)
1217.295 Million cell updates/sec

Title: SEQ4-THEN-SEQ3

Perfect score: 40

Sequence: 1 atccagattcttggcggtagagagctgttggtcagactt 40

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 0.1

Searched: 4966997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

- 1: Geneseqn1990s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	24.6	61.5	456	10 ADB50561	Adb50561 Primary r
C 2	24.6	61.5	1105	10 ADH54859	Adh54859 Rat insul
C 3	24.6	61.5	1640	10 ADH54925	Adh54925 Rat insul
C 4	23.8	59.5	860	6 ABK36163	Abk36163 cDNA sequ
C 5	23.8	59.5	871	4 AAH05357	Aah05357 Human cDN
C 6	23.8	59.5	1928	6 ABK36179	Abk36179 cDNA sequ
C 7	23.8	59.5	2343	4 AAH16123	Aah16123 Human cDN
C 8	23.8	59.5	2390	4 AAH14062	Aah14062 Human cDN
C 9	23.8	59.5	2650	3 AAZ93150	Aaz93150 Sequence
C 10	23.8	59.5	2770	3 AAC93410	Aac93410 Human sec
C 11	23.8	59.5	2892	4 AAH17602	Aah17602 Human cDN
C 12	23.8	59.5	3134	6 ABQ54736	Abq54736 Human ova
C 13	23.8	59.5	6701	4 AAH16661	Aah16661 Human cDN
C 14	23.8	59.5	7228	8 ACAS6981	Acas6981 Human adi
C 15	23.8	59.5	7747	5 AAS44988	Aas44988 cDNA enco
C 16	23.8	59.5	7827	4 AAI58227	Aai58227 Human pol
C 17	23.8	59.5	7827	5 ADQ98434	Adq98434 DNA encod
C 18	23.8	59.5	7827	9 ADB48194	Adb48194 Novel hum
C 19	23.8	59.5	8825	10 ADF81716	Adf81716 Leukaemia

C 20	23.8	59.5	9508	10 ADE40129	Ade40129 Human NOV
C 21	23.8	59.5	15552	13 ADR83456	Adr83456 Human ret
C 22	23.8	59.5	15552	14 ADX06153	Adx06153 Cyclin-de
C 23	23.6	59.0	2157	13 ADS45710	Ads45710 Bacterial
C 24	23	57.5	1266	9 ADB07057	Adb07057 Alloiooc
C 25	23	57.5	1266	9 ADB07055	Adb07055 Alloiooc
C 26	23	57.5	110000	9 ADB12064_02	Continuation (3 of
C 27	22.6	56.5	1107	5 AAH66350	Aah66350 C glutami
C 28	22.6	56.5	1164	8 ACA00491	Aca00491 C. glutam
C 29	22.6	56.5	1230	14 AEB15386	Aeb15386 C glutam
C 30	22.6	56.5	1294	10 ADD13935	Add13935 C. glutam
C 31	22.6	56.5	1600	6 AAL46365	Aal46365 C glutami
C 32	22.6	56.5	34980	5 AAH68528	Aah68528 C glutami
C 33	22.2	55.5	210	10 AAD59410	Aad59410 Mouse Kvb
C 34	22.2	55.5	262	4 ABA75082	Aba75082 Human foe
C 35	22.2	55.5	262	4 AAI55616	Aai55616 Probe #24
C 36	22.2	55.5	262	4 AAk23597	Aak23597 Human bra
C 37	22.2	55.5	262	4 ABS49351	Abs49351 Human liv
C 38	22.2	55.5	262	6 ABS23215	Abs23215 Human liv
C 39	22.2	55.5	572	4 ABA62561	Aba62561 Human foe
C 40	22.2	55.5	572	4 AAI42547	Aai42547 Probe #11
C 41	22.2	55.5	572	4 AAk10917	Aak10917 Human bra
C 42	22.2	55.5	572	4 ABS36435	Abs36435 Human liv
C 43	22.2	55.5	572	6 ABS10777	Abs10777 Human gen
C 44	22.2	55.5	1706	14 ADX25701	Adx25701 Novel cel
C 45	22.2	55.5	1706	14 ADX26166	Adx26166 Novel cel

ALIGNMENTS

RESULT 1

ADB50561/c

ID ADB50561 standard; DNA; 456 BP.

XX ADB50561;

DT 04-DEC-2003 (first entry)

XX Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:1103.

DE toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;

XX toxicity marker; toxicity progression; drug screening;

XX primary rat hepatocyte toxicity modelling; gene; ds.

OS Rattus norvegicus.

XX WO2003065993-A2.

XX 14-AUG-2003.

XX 04-FEB-2003; 2003WO-US003482.

XX 04-FEB-2002; 2002US-0353171P.

XX 03-MAR-2002; 2002US-0363534P.

XX 08-APR-2002; 2002US-0370248P.

XX 10-APR-2002; 2002US-0371134P.

XX 10-APR-2002; 2002US-0371135P.

XX 10-APR-2002; 2002US-0371150P.

XX 11-APR-2002; 2002US-0371413P.

XX 19-APR-2002; 2002US-0373601P.

XX 19-APR-2002; 2002US-0373602P.

XX 22-APR-2002; 2002US-0374139P.

XX 08-MAY-2002; 2002US-0378370P.

XX 09-MAY-2002; 2002US-0378652P.

XX 09-MAY-2002; 2002US-0378653P.

XX 09-MAY-2002; 2002US-0378655P.

XX 09-JUL-2002; 2002US-0394230P.

XX 09-JUL-2002; 2002US-0394253P.

XX 04-SEP-2002; 2002US-0407688P.

XX 28-JAN-2003; 2003US-0442900P.

XX (GENE-) GENE LOGIC INC.

PA

```

XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Blashoff M;
XX WPI; 2003-731472/69.
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
XX Claim 44; SEQ ID NO 1103; 874pp; English.
XX
XX The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
XX Sequence 456 BP; 123 A; 95 C; 103 G; 135 T; 0 U; 0 Other;
SQ
Query Match 61.5%; Score 24.6; DB 10; Length 456;
Best Local Similarity 76.9%; Pred. No. 4.6;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATCCAGATTCTTGGCGGTAGACAGCTGCTTGTCAGACT 39
Db 369 ATCCTGATTCTTGGCGGTAGACACACAGCTTGCCACACT 331
RESULT 2
ADH54859
ID ADH54859 standard; cDNA; 1105 BP.
XX
XX ADH54859;
XX
XX 25-MAR-2004 (first entry)
XX
XX Rat insulinoma differentially expressed clone 8.
XX
XX Rat; insulinoma differentially expressed clone; ss; glucose response;
KW insulin; diabetes.
XX
XX Rattus sp.
XX
XX US2003148421-A1.
XX
XX 07-AUG-2003.
XX
XX 19-FEB-2002; 2002US-00080381.
XX
XX 20-FEB-2001; 2001US-0270251P.
XX
XX 09-MAR-2001; 2001US-0274706P.
XX
XX 15-MAY-2001; 2001US-0291354P.
XX
XX (NEWG/) NEWGARD C B.
XX
XX (JENS/) JENSEN P B.
XX
XX Newgard CB, Jensen PB;
XX
XX WPI; 2003-897569/82.
XX
XX New polynucleotide, useful for enhancing, stabilizing or introducing
PT glucose-responsiveness in a host cell that secretes insulin, or as target

```

```

PT for drug screening or diagnostic indicator for the loss of glucose-
XX responsiveness.
XX
XX Claim 1; SEQ ID NO 1; 93pp; English.
XX
XX The invention relates to an isolated polynucleotide (differentially
CC expressed by insulinoma cells) encoding a protein involved in regulating
CC glucose responses. Also included are an expression cassette (comprising
CC the polynucleotide and a promoter, where the polynucleotide is positioned
CC so as to be under the regulatory control of the promoter), an
CC oligonucleotide of 15-50 bases (comprising 15 or more bases of
CC polynucleotide sequence), a cell comprising the heterologous
CC polynucleotide encoding the polypeptide, a method of expressing a
CC polypeptide, a polypeptide comprising a sequence encoded by the
CC polynucleotide, a peptide of 5-35 residues comprising 5 or more
CC consecutive residues of the polypeptide, a monoclonal antibody that binds
CC immunologically to the polypeptide, a nucleic acid encoding the antibody,
CC a polyclonal antibody composition, where the antibodies bind
CC immunologically to the polypeptide, a method for modulating the glucose-
CC responsiveness of a cell, a method of screening for a modulator of
CC expression of the polypeptide and a method for measuring the expression
CC of the polypeptide. The polynucleotide is useful for enhancing, the
CC stabilising or introducing glucose-responsiveness in a host cell that
CC secretes insulin, and as target for drug screening or diagnostic indicator
CC for the loss of glucose-responsiveness e.g. in diabetes. The present
CC sequence is an isolated polynucleotide (differentially expressed by
CC insulinoma cells) encoding a protein involved in regulating glucose
CC responses of the invention.
XX
XX Sequence 1105 BP; 286 A; 266 C; 255 G; 297 T; 0 U; 1 Other;
SQ
Query Match 61.5%; Score 24.6; DB 10; Length 1105;
Best Local Similarity 76.9%; Pred. No. 5.5;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATCCAGATTCTTGGCGGTAGACAGCTGCTTGTCAGACT 39
Db 742 ATCCTGATTCTTGGCGGTAGACACACAGCTTGCCACACT 780
RESULT 3
ADH54925
ID ADH54925 standard; cDNA; 1640 BP.
XX
XX ADH54925;
XX
XX 25-MAR-2004 (first entry)
XX
XX Rat insulinoma differentially expressed clone 8 3' sequence.
XX
XX Rat; insulinoma differentially expressed clone; ss; glucose response;
KW insulin; diabetes.
XX
XX Rattus sp.
XX
XX US2003148421-A1.
XX
XX 07-AUG-2003.
XX
XX 19-FEB-2002; 2002US-00080381.
XX
XX 20-FEB-2001; 2001US-0270251P.
XX
XX 09-MAR-2001; 2001US-0274706P.
XX
XX 15-MAY-2001; 2001US-0291354P.
XX
XX (NEWG/) NEWGARD C B.
XX
XX (JENS/) JENSEN P B.
XX
XX Newgard CB, Jensen PB;
XX
XX WPI; 2003-897569/82.
XX
XX New polynucleotide, useful for enhancing, stabilizing or introducing
PT glucose-responsiveness in a host cell that secretes insulin, or as target

```

PT glucose-responsiveness in a host cell that secretes insulin, or as target
PT for drug screening or diagnostic indicator for the loss of glucose-
XX responsiveness.
PS Claim 1; SEQ ID NO 67; 93pp; English.
XX
CC The invention relates to an isolated polynucleotide (differentially
CC expressed by Insulinoma cells) encoding a protein involved in regulating
CC glucose responses. Also included are an expression cassette (comprising
CC the polynucleotide and a promoter, where the polynucleotide is positioned
CC so as to be under the regulatory control of the promoter), an
CC oligonucleotide of 15-50 bases (comprising 15 or more bases of
CC polynucleotide sequence), a cell comprising the heterologous
CC polynucleotide encoding the polypeptide, a method of expressing a
CC polypeptide, a polypeptide comprising a sequence encoded by the
CC polynucleotide, a peptide of 5-35 residues comprising 5 or more
CC consecutive residues of the polypeptide, a monoclonal antibody that binds
CC immunologically to the polypeptide, a nucleic acid encoding the antibody,
CC a polyclonal antibody composition, where the antibodies bind
CC immunologically to the polypeptide, a method for modulating the glucose-
CC responsiveness of a cell, a method of screening for a modulator of
CC expression of the polypeptide and a method for measuring the expression
CC of the polypeptide. The polynucleotide is useful for enhancing,
CC stabilising or introducing glucose-responsiveness in a host cell that
CC secretes insulin, and as target for drug screening or diagnostic indicator
CC for the loss of glucose-responsiveness e.g. in diabetes. The present
CC sequence is a 5' or 3' sequence fragment of an isolated polynucleotide
CC (differentially expressed by Insulinoma cells) encoding a protein
CC involved in regulating glucose responses.
XX
XX Sequence 1640 BP; 388 A; 404 C; 365 G; 472 T; 0 U; 11 Other;
SQ
Query Match 61.5%; Score 24.6; DB 10; Length 1640;
Best Local Similarity 76.9%; Pred. NO. 6;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATCCGATCTTGGCGGTAGAGCTGCTTGTCACACT 39
|||||
DB 1277 ATCCTGATCTTGGCGGTAGACACACAGCTTGCCACACT 1315

RESULT 4
ABK36163/c
ID ABK36163 standard; cDNA; 860 BP.
XX
AC ABK36163;
XX
DT 08-MAY-2002 (first entry)
XX
DE cDNA sequence #554 encoding novel human secreted protein.
XX
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW immune deficiency disorder; blood disorder; inflammatory disorder;
KW infectious disorder; allergic condition; neurodegenerative disorder;
KW liver fibrosis; coagulation disorder; gene therapy; antimicrobial;
KW tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200177289-A2.
PN
XX 18-OCT-2001.
PD
XX 29-MAR-2001; 2001WO-US010232.
PP
XX 06-APR-2000; 2000US-0195605P.
PR
XX (GEM) GENETICS INST INC.
PA
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
XX Merberg D, Tracy M, Acostino MJ, Bowman MR, Spaulding V, Wong GG;
XX Clark HF, Fechtel K, Howes SH, Resnick RJ, Gulukota K, Graham JR;
XX

DR WPI; 2002-179322/23.
XX Six hundred and twenty three polynucleotides derived from a variety of
PT human tissue sources which encode secreted proteins, useful for treating
PT immune deficiencies and disorders such as autoimmune disorders.
XX
PS Claim 1; Page 362; 393pp; English.
XX
CC The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides a
CC method for producing proteins from these polynucleotide sequences. The
CC proteins are useful for identifying compounds that modulate their
CC activity and production. The sequences of the invention are useful for
CC treating diseases such as hyperproliferative disorders (e.g. cancer),
CC immune deficiency disorders (e.g. severe combined immunodeficiency
CC (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
CC (e.g. thrombocytopaenia), inflammatory disorders (e.g. arthritis),
CC infectious disorders (e.g. hepatitis), allergic conditions (e.g. asthma),
CC neurodegenerative disorders (e.g. Alzheimer's disease), liver fibrosis,
CC coagulation disorders (e.g. haemophilia), and tumours. The polynucleotide
CC sequences of the invention are also useful in gene therapy. ABK3610-
CC ABK36232 represent the cDNA sequences of the invention that encode for
CC novel human secreted proteins
XX
XX Sequence 860 BP; 225 A; 236 C; 240 G; 159 T; 0 U; 0 Other;
SQ
Query Match 59.5%; Score 23.8; DB 6; Length 860;
Best Local Similarity 80.0%; Pred. NO. 11;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCGATCTTGGCGGTAGAGCTGCTTGCTCA 35
|||||
DB 55 ATCCAAATCTTGGCGCTAGGATGCTTTTCA 21

RESULT 5
AAH05357/c
ID AAH05357 standard; cDNA; 871 BP.
XX
AC AAH05357;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:2192.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
XX EP1074617-A2.
PN
XX 07-FEB-2001.
PP
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.

CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 CC
 SQ Sequence 2343 BP; 569 A; 643 C; 625 G; 506 T; 0 U; 0 Other;

Query Match 59.5%; Score 23.8; DB 4; Length 2343;
 Best Local Similarity 80.0%; Pred. No. 14;
 Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ATCCAGATCTTCGGCGGTAGAGAGCTGCTTGCTCA 35
 |||||
 Db 591 ATCCAAATCTTGGCGCTAGGATGCTTTTCA 557

RESULT 8
 AAH14062/c
 ID AAH14062 standard; cDNA; 2390 BP.
 XX
 AC AAH14062;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:11199.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 FT Primer sets for synthesising polynucleotides, particularly the 5602 full-
 FT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 11199; 2537pp + Sequence Listing; English.
 XX

CC The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 CC
 SQ Sequence 2390 BP; 580 A; 656 C; 639 G; 515 T; 0 U; 0 Other;
 Query Match 59.5%; Score 23.8; DB 4; Length 2390;
 Best Local Similarity 80.0%; Pred. No. 14;
 Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ATCCAGATCTTCGGCGGTAGAGAGCTGCTTGCTCA 35
 |||||
 Db 649 ATCCAAATCTTGGCGCTAGGATGCTTTTCA 615

RESULT 9
 AAZ93150/c
 ID AAZ93150 standard; DNA; 2650 BP.
 XX
 AC AAZ93150;
 XX
 DT 19-JUN-2000 (first entry)
 XX
 DE Sequence encoding DNA replication and repair associated protein.
 XX
 KW DRASP; replication; repair; treatment; prophylaxis; diagnosis; screening;
 KW anemia; epilepsy; hypothyroidism; cancer; autoimmune disease; AIDS;
 KW acquired immune deficiency syndrome; atherosclerosis;
 KW autoimmune thyroiditis; bronchitis; allergy; asthma; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 77..2317
 FT /*tag= a
 FT /product= "DNA replication and repair associated protein"
 XX
 PN W0200008156-A2.
 XX
 PD 17-FEB-2000.
 XX
 PF 06-AUG-1999; 99WO-US017800.
 XX
 PR 07-AUG-1998; 98US-0155245P.
 PR 27-AUG-1998; 98US-0155181P.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Tang YT, Corley NC, Baughn MR, Reddy R, Guegler KJ, Yue H;
 XX WPI; 2000-205711/18.
 DR P-PSDB; AAY82742.
 XX
 FT New proteins useful for diagnosing, treating or preventing disorders

PT associated with expression of DNA replication- and repair- associated
PT proteins.
PS Claim 9; Page 68-69; 70pp; English.
XX
CC DNA replication and repair associated proteins (DRASP) and their
CC antagonists are useful for treating or preventing a disease associated
CC with increased expression or activity of DRASP. These disorders include
CC development disorders e.g anemia, epilepsy, hypothyroidism, proliferative
CC disorders including cancer; autoimmune/inflammatory disorders including
CC AIDS, atherosclerosis; autoimmune thyroiditis; bronchitis; allergies and
CC asthma. The DRASP's and the polynucleotides encoding them are also useful
CC for diagnosis of the above mentioned disorders. Catalytic or immunogenic
CC fragments of the DRASP or its oligopeptides can be useful for screening
CC libraries of compounds in a variety of drug screening techniques
XX
SQ Sequence 2650 BP; 667 A; 697 C; 710 G; 576 T; 0 U; 0 Other;
Query Match 59.5%; Score 23.8; DB 3; Length 2650;
Best Local Similarity 80.0%; Pred. No. 14;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATCCGATTCTTGGCGGTAGAGAGCTGCTTGCTCA 35
Db 883 ATCCAAATTTCTGGCGCTAGGGATGCTTTTCA 849
RESULT 10
AAC93410/c
ID AAC93410 standard; cDNA; 2770 BP.
AC AAC93410;
XX
DT 26-FEB-2001 (first entry)
XX
DE Human secreted protein gene 47 SEQ ID NO:57.
XX
KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW cerebrovascular disorder; nervous system disorder; ocular disorder;
KW wound healing; skin aging; food additive; preservative; ss.
XX
OS Homo sapiens.
XX
XX WO200061620-A1.
PN
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009069.
XX
XX 09-APR-1999; 99US-0128702P.
PR
XX 20-JAN-2000; 2000US-0177049P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX (ROSE/) ROSEN C A.
PA
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
PI WPI; 2000-619225/59.
DR
XX P-PADB; AAB51666.
DR
XX Isolated nucleic acid molecule encoding a human secreted protein is used
PT in preventing, treating or ameliorating a medical condition.
XX
PS Claim 1; Page 461-462; 540pp; English.
XX
XX The polynucleotide sequences given in AAC93364 to AAC93412 encode the
CC human secreted proteins given in AAB51620 to AAB51668. AAB51669 to
CC AAB51722 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present

CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
CC cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and polypeptides can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders, cardiovascular
CC disorders, cerebrovascular disorders, angiogenesis, nervous system
CC disorders, infections caused by bacteria, viruses and fungi and ocular
CC disorders. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. AAC93355 to AAC93363 and
CC AAB51619 represent sequences which are used in the exemplification of the
CC present invention
XX
SQ Sequence 2770 BP; 703 A; 721 C; 746 G; 598 T; 0 U; 2 Other;
Query Match 59.5%; Score 23.8; DB 3; Length 2770;
Best Local Similarity 80.0%; Pred. No. 14;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATCCGATTCTTGGCGGTAGAGAGCTGCTTGCTCA 35
Db 1001 ATCCAAATTTCTGGCGCTAGGGATGCTTTTCA 967
RESULT 11
AAH17602/c
ID AAH17602 standard; cDNA; 2892 BP.
XX
AC AAH17602;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:17110.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX EP1074617-A2.
PN
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
PR
XX 27-AUG-1999; 99JP-00300253.
PR
XX 11-JAN-2000; 2000JP-00118776.
PR
XX 02-MAY-2000; 2000JP-00183767.
PR
XX 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
DR
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 8; SEQ ID NO 17110; 2537pp + Sequence Listing; English.
PS
XX

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 2892 BP; 721 A; 756 C; 785 G; 630 T; 0 U; 0 Other;

Query Match 59.5%; Score 23.8; DB 4; Length 2892;
 Best Local Similarity 80.0%; Pred. No. 14;
 Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATTCCTGGCGCTAGAGAGCTGCTGGTCA 35
 DB 1144 ATCCAAATTCCTGGCGCTAGGATGCTTTTCA 1110

RESULT 12

ID ABQ54736/c
 AC ABQ54736
 XX
 XX
 XX
 XX
 XX
 XX

Human ovarian antigen HOCFY47 cDNA, SEQ ID NO:616.

Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; gene; ss.

OS Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX P-PSDB; ABP41659.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,

PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.

PS Claim 1; SEQ ID NO 616; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 3134 BP; 789 A; 801 C; 855 G; 678 T; 0 U; 11 Other;

Query Match 59.5%; Score 23.8; DB 6; Length 3134;

Best Local Similarity 80.0%; Pred. No. 15;

Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATTCCTGGCGCTAGAGAGCTGCTGGTCA 35
 DB 1330 ATCCAAATTCCTGGCGCTAGGATGCTTTTCA 1296

RESULT 13

AAH16661/c

ID AAH16661 standard; cDNA; 6701 BP.

XX AAH16661;

AC AAH16661;

DT 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:15799.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

PA (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
XX Claim 8; SEQ ID NO 15799; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX
SQ Sequence 6701 BP; 1676 A; 1812 C; 1722 G; 1491 T; 0 U; 0 Other;
Query Match 59.5%; Score 23.8; DB 4; Length 6701;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATCCAGATTCCTGGCGGTAGAGAGCTGTGGTCA 35
DB 4953 ATCCAAATTCCTGGCGGTAGAGATGCTTTTCA 4919
RESULT 14
ACA56981/c
ID ACA56981 standard; cDNA; 7228 BP.
XX
XX ACA56981;
XX
XX 10-JUN-2003 (first entry)
XX
XX Human adipocyte Selected Interacting domain, SID, cDNA #68.
XX
XX Human; ss; gene; prey; adipocyte; SID; selected interacting domain; anorectic; antidiabetic; protein-protein interaction; diabetes;
XX yeast 2-hybrid assay; metabolic disorder; obesity.
XX
XX Homo sapiens.
XX
XX WO200286122-A2.
XX
XX 31-OCT-2002.
XX
XX 14-MAR-2002; 2002WO-EP003768.
XX
XX 14-MAR-2001; 2001US-0275734P.
XX
XX (HYBR-) HYBRIGENICS.
PA

XX
PI Legrain P, Daviet L;
XX WPI; 2003-103412/09.
DR P-PSDB; ABU70437.
XX
XX New complex between two interacting proteins in adipocyte cells, useful for identifying selected interacting domains that modulate protein interactions, or for preventing or treating metabolic disorders such as obesity or diabetes.
XX
XX Claim 7; Page 134-138; 382pp; English.
XX
XX The invention relates to a complex between two interacting proteins in adipocyte cells, given in the specification. The proteins are identified by selecting a bait protein from a known adipocyte marker and then performing a yeast 2-hybrid selection to isolate prey proteins encoded by members of an adipocyte cDNA library. The proteins are designated SID (RTM) (selected interacting domains) proteins. Also included are a polynucleotide encoding a polypeptide in the adipocyte cells, a recombinant host cell expressing at least one of the interacting polypeptides of the complex, selecting a modulating compound in adipocyte cells, a SID (RTM) polypeptide comprising any of the 738 amino acid sequences given in the specification (including its fragment or variant), a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences given in the specification (including its fragment or variant), a vector comprising the SID (RTM) polynucleotide, a recombinant host cell comprising the vector, a protein chip comprising the polypeptides and a record comprising all or part of the data, listed in the specification. The complex, polypeptides, polynucleotides and compounds are useful for preventing or treating metabolic disorders such as obesity or diabetes. The polynucleotides are useful as probes or primers. The complex is particularly useful for identifying selected interacting domains (SID (RTM)) for screening drugs that modulate the protein interaction, thus exhibiting the therapeutic effect. The present sequence encodes a SID (prey) protein of the invention
XX
SQ Sequence 7228 BP; 1822 A; 1955 C; 1903 G; 1548 T; 0 U; 0 Other;
Query Match 59.5%; Score 23.8; DB 8; Length 7228;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATCCAGATTCCTGGCGGTAGAGAGCTGTGGTCA 35
DB 5794 ATCCAAATTCCTGGCGGTAGAGATGCTTTTCA 5760
RESULT 15
AAS44988/c
ID AAS44988 standard; cDNA; 7747 BP.
XX
XX AAS44988;
XX
XX 18-DEC-2001 (first entry)
XX
XX cDNA encoding novel human secretory protein, Seq ID No 69.
XX
XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock; ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
XX transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
XX amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
XX ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
XX gut protection; lung; liver fibrosis; immune deficiency; infection;
XX severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
XX multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
XX fertility; analgesic; pain; antigen; ss.
XX
XX Homo sapiens.
XX
XX WO200166689-A2.
XX
XX 13-SEP-2001.
PD

XX
PF 05-MAR-2001; 2001WO-US004942.
XX
XX 07-MAR-2000; 2000US-00519705.
PR 19-MAY-2000; 2000US-00574454.
PR 17-JUN-2000; 2000US-00596193.
PR 14-JUL-2000; 2000US-00616847.
PR 19-SEP-2000; 2000US-00665363.
PR 20-OCT-2000; 2000US-00693267.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
DR P-PSDB; AAU28088.
DR
XX
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders.
XX
XX
PS Claim 1; SEQ ID NO 69; 107pp; English.
XX
XX The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,
CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,
CC or periodontal disease. Furthermore, (I) is also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAS44920-AAS45295 represent novel human secreted protein
CC coding sequences of the invention
XX
SQ Sequence 7747 BP; 1948 A; 2083 C; 2034 G; 1682 T; 0 U; 0 Other;

Query Match 59.5%; Score 23.8; DB 5; Length 7747;
Best Local Similarity 80.0%; Pred. No. 18;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATTCCTGGCGTAGAGAGCTGCTGGTCA 35
|||||
DB 5983 ATCCAAATTCCTGGCGTAGAGATGCTTTTCA 5949
|||||

Search completed: April 18, 2006, 19:13:01
Job time : 222 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:13:18 ; Search time 1706.5 Seconds
(without alignments)
1096.679 Million cell updates/sec

Title: SEQ4-THEN-SEQ3

Perfect score: 40

Sequence: 1 atccagattcttggcggttagagagctgtgttcagactt 40

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 0.1

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hic:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_gss1:*

10: gb_gss2:*

11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	25.4	63.5	535	2	BE682696
C 2	24.8	62.0	748	2	RG391395
C 3	24.6	61.5	456	1	AA957723
C 4	24.6	61.5	484	2	BF420664
C 5	24.6	61.5	502	2	BF557216
C 6	24.6	61.5	643	7	CK843614
C 7	23.8	59.5	308	1	AA143624
C 8	23.8	59.5	349	8	N59506
C 9	23.8	59.5	355	2	RG952796
C 10	23.8	59.5	355	3	BQ318970
C 11	23.8	59.5	358	1	AW479483
C 12	23.8	59.5	367	2	BF767919
C 13	23.8	59.5	436	1	AA159119
C 14	23.8	59.5	443	6	CD673609
C 15	23.8	59.5	471	2	BE485052
C 16	23.8	59.5	482	8	DN601950
C 17	23.8	59.5	501	1	AW579072
C 18	23.8	59.5	546	3	BM712229
C 19	23.8	59.5	550	3	BP264713
C 20	23.8	59.5	568	6	CB150078
C 21	23.8	59.5	580	3	BP196040
C 22	23.8	59.5	581	3	BP344471

C 23	23.8	59.5	583	3	BP227255
C 24	23.8	59.5	583	3	BP255553
C 25	23.8	59.5	583	3	BP345331
C 26	23.8	59.5	588	2	BE619266
C 27	23.8	59.5	641	3	BM384410
C 28	23.8	59.5	641	3	BM833752
C 29	23.8	59.5	672	2	BG749832
C 30	23.8	59.5	678	2	BE905313
C 31	23.8	59.5	692	3	BM724099
C 32	23.8	59.5	706	2	EG388908
C 33	23.8	59.5	715	3	BI916606
C 34	23.8	59.5	715	7	CR992502
C 35	23.8	59.5	726	2	BG764111
C 36	23.8	59.5	735	7	CT000492
C 37	23.8	59.5	739	7	CR980220
C 38	23.8	59.5	762	3	BP223482
C 39	23.8	59.5	764	1	AU132166
C 40	23.8	59.5	815	2	BG753212
C 41	23.8	59.5	818	3	BI913289
C 42	23.8	59.5	832	2	BP338327
C 43	23.8	59.5	840	2	EG536386
C 44	23.8	59.5	869	1	AU124415
C 45	23.8	59.5	871	1	AU124249

ALIGNMENTS

RESULT 1
LOCUS BE682696/c 535 bp mRNA linear EST 25-APR-2001
DEFINITION 180928 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE682696
VERSION BE682696.1 GI:10068819
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 535)
AUTHORS Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Perlea, G., Holt, I., Karanycheva, S., Liang, F., Quackenbush, J. and Keefe, J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
PUBMED 11282978
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACACGCTATGACCAT
REVERSE: GTTTCGCCAGTCACGACG
Plate: 75 row: C column: 23
Seq primer: ATTAGGTGACACTAG.
Location/Qualifiers
1. .535
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4BOV"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;

BP227255	BP227255
BP255553	BP255553
BP345331	BP345331
BE619266	601473137
BE384410	601277738
BM833752	K-EST0108
EG749832	602708221
BE905313	601499172
BM724099	UI-E-B01-
EG388908	602414648
BI916606	603178537
CR992502	CR992502
BG764111	602737128
CT000492	CT000492
CR980220	CR980220
BP223482	BP223482
AU132166	AU132166
BG753212	602731604
BI913289	603180045
BP338327	602035559
EG536386	602564779
AU124415	AU124415
AU124249	AU124249

Library made from pooled tissue from day 20 and day 40 embryos."

ORIGIN
Query Match 63.5%; Score 25.4; DB 2; Length 535;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ATCCAGATTCTGGCGGTAGAGCTGCTTGGTCA 35
|||||
Db 294 ATCCAGATTCTGGCGGTGGGATGCTTCTTCA 260
|||||

RESULT 2
BG391395/c
LOCUS BG391395 748 bp mRNA linear EST 12-MAR-2001
DEFINITION 602417472F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4537192 5',
mRNA sequence.

ACCESSION BG391395 GI:13284843
VERSION BG391395
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 748)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10461 row: o column: 17
High quality sequence stop: 734.

FEATURES
source
1..748
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4537192"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 62.0%; Score 24.8; DB 2; Length 748;
Best Local Similarity 80.6%; Pred. No. 57;
Matches 29; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATTCTGGCGGTAGAGCTGCTTGGTCA 36
|||||
Db 585 ATCCAAATTCTTGGCGGTAGGATGTCTTTTCAG 550
|||||

RESULT 3
AA957723/c
LOCUS AA957723 456 bp mRNA linear EST 04-JUL-1999
DEFINITION UI-R-E1-gc-a-12-0-UI_s1 UI-R-E1 Rattus norvegicus cDNA clone
UI-R-E1-gc-a-12-0-UI 3', mRNA sequence.
ACCESSION AA957723
VERSION AA957723.1 GI:4277613

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE

JOURNAL
PUBMED
COMMENT

EST.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus;
1 (bases 1 to 456)
Bonaldo, M.P., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
8889548
On May 7, 1998 this sequence version replaced gi:3121418.
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence tag present in the cDNA between the NotI site and the
oligo-dt track served to identify it as a clone from the normalized
adult 12-day-embryo library. cDNA Library Preparation: M. Fatima
Bonaldo, Ph.D. Clone distribution: clones will be available through
Research Genetics This clone is also available through the
I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov). IMAGE
ID=1772375 The following repetitive elements were found in this
cDNA sequence: 140-277, >RSINE1#SINE/B4(B5 301-456,
>MER31B#LTR/MER4-group
Seq primer: M13 Forward
POLYA=No.

FEATURES
source
Location/Qualifiers
1..456
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-E1-gc-a-12-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-E1"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-E1
library is a subtracted library derived from the UI-R-E0
library. The UI-R-E0 library consisted of a mixture of
individually tagged normalized libraries constructed from
8, 12 and 18-day embryo. The tag is a string of 3-5
nucleotides present between the Not I site and the
oligo-dt track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-E1) was constructed as follows: PCR
amplified cDNA inserts from a pool of UI-R-E0 clones from
which 3, ESTs had been derived was used as a driver in a
hybridization with the UI-R-E0 library in the form of a
single-stranded circles. The remaining single-stranded
circles (subtracted library) was purified by
hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the UI-R-E1
library. This procedure has been previously described
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
1996)"

Query Match 61.5%; Score 24.6; DB 1; Length 456;
Best Local Similarity 76.9%; Pred. No. 64;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATCCAGATTCTGGCGGTAGAGCTGCTTGGTCACT 39
|||||
Db 369 ATCTCGATTCTTGGCGGTAGAGCAACAGCTTGGCCCACT 331
|||||

```

RESULT 4
BP420664/c
LOCUS
DEFINITION
  BP420664
  UI-R-BJ2-bpy-f-09-0-UI.s1 484 bp mRNA linear EST 28-NOV-2000
  UI-R-BJ2-bpy-f-09-0-UI 3', mRNA sequence.
ACCESSION
  BP420664
  BP420664.1 GI:11408653
KEYWORDS
  EST.
SOURCE
  Rattus norvegicus (Norway rat)
ORGANISM
  Rattus norvegicus
  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
  Sciurognathi; Muroidae; Muridae; Murinae; Rattus.
REFERENCE
  1 (bases 1 to 484)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
JOURNAL
  PUBMED
  8889548
COMMENT
  Contact: Soares, MB
  Coordinated Laboratory for Computational Genomics
  University of Iowa
  375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: bento-soares@uiowa.edu
  The sequence contained an oligo-dT track that was present in the
  oligonucleotide that was used to prime the synthesis of first
  strand cDNA and therefore this may represent a bonafide poly A
  tail. The sequence tag present in the cDNA between the NotI site
  and the oligo-dT track served to identify it as a clone from the
  normalized atrium at 16.5 dpc library cDNA Library Preparation:
  M.B. Soares Lab Clone distribution: Clones will be available
  through Research Genetics (www.resgen.com) The following repetitive
  elements were found in this cDNA sequence: 141-278,
  >RSINE1$SINE/B4(B5 302-483, >MER31B$LTR/MER4-group
  Seq primer: M13 Forward
  POLYA=Yes.
FEATURES
  source
  1. .484
  /location/Qualifiers
  /organism="Rattus norvegicus"
  /mol_type="mRNA"
  /strain="Sprague-Dawley"
  /db_xref="taxon:10116"
  /clone="UI-R-BJ2-bpy-f-09-0-UI"
  /lab_host="DH10B (Life Technologies)"
  /clone_lib="UI-R-BJ2"
  /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
  polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ2
  library is a subtracted library derived from the following
  tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc,
  atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15
  dpc, AV canal at 15 dpc. For a detailed description of
  the library from which this clone was derived, please
  visit our web site at ratest.eng.uiowa.edu. The
  subtraction has been previously described in (Bonaldo,
  Lennon and Soares, Genome Research 6:791-806, 1996)
  TAG TISSUE=atrium at 16.5 dpc
  TAG LIB=UI-R-BJ2
  TAG_SEQ=GATTC"
ORIGIN
  Query Match 61.5%; Score 24.6; DB 2; Length 484;
  Best Local Similarity 76.9%; Pred. No. 64;
  Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
  QY 1 ATCCAGATTTCTGGCGGTAGAGAGCTGCTTGTCAGACT 39
  |||||
  DB 370 ATCTGTGATTTCTGGCGGTAGACACAGCTTGCCCACT 332
  |||||

RESULT 5
BP557216
LOCUS
DEFINITION
  BP557216
  UI-R-E1-gc-a-12-0-UI.r1 502 bp mRNA linear EST 12-DEC-2000
  UI-R-E1-gc-a-12-0-UI 5', mRNA sequence.
ACCESSION
  BP557216
  BP557216.1 GI:11666946
KEYWORDS
  EST.
SOURCE
  Rattus norvegicus (Norway rat)
ORGANISM
  Rattus norvegicus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
  Sciurognathi; Muroidae; Muridae; Murinae; Rattus.
REFERENCE
  1 (bases 1 to 502)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
JOURNAL
  PUBMED
  8889548
COMMENT
  Contact: Soares, MB
  Coordinated Laboratory for Computational Genomics
  University of Iowa
  375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: bento-soares@uiowa.edu
  cDNA Library Preparation: M.B. Soares Lab Clone distribution:
  clones will be available through Research Genetics (www.resgen.com)
  This clone is also available through the I.M.A.G.E. Consortium at
  LLNL (infoimage.llnl.gov). IMAGE ID= 1772375 The following
  repetitive elements were found in this cDNA sequence: 253-472,
  >MER31B$LTR/MER4-group
  Seq primer: M13 Forward.
FEATURES
  source
  1. .502
  /location/Qualifiers
  /organism="Rattus norvegicus"
  /mol_type="mRNA"
  /strain="Sprague-Dawley"
  /db_xref="taxon:10116"
  /clone="UI-R-E1-gc-a-12-0-UI"
  /dev_stage="adult"
  /lab_host="DH10B (Life Technologies)"
  /clone_lib="UI-R-E1"
  /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
  polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-E1
  library is a subtracted library derived from the UI-R-E0
  library. The UI-R-E0 library consisted of a mixture of
  individually tagged normalized libraries constructed from
  8, 12 and 18-day embryo. The tag is a string of 3-5
  nucleotides present between the Not I site and the
  oligo-dT track which allows identification of the library
  of origin of a clone within the mixture. The subtracted
  library (UI-R-E1) was constructed as follows: PCR
  amplified cDNA inserts from a pool of UI-R-E0 clones from
  which 3' ESTs had been derived was used as a driver in a
  hybridization with the UI-R-E0 library in the form of
  single-stranded circles. The remaining single-stranded
  circles (subtracted library) was purified by
  hydroxyapatite column chromatography, converted to
  double-stranded circles and electroporated into DH10B
  bacteria (Life Technologies) to generate the UI-R-E1
  library. This procedure has been previously described
  (Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
  1996)"
ORIGIN
  Query Match 61.5%; Score 24.6; DB 2; Length 502;
  Best Local Similarity 76.9%; Pred. No. 65;
  Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
  QY 1 ATCCAGATTTCTTGGCGGTAGAGAGCTGCTTGTCAGACT 39
  |||||
  DB 422 ATCTGTGATTTCTGGCGGTAGACACAGCTTGCCCACT 460
  |||||

```


Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

TITLE JOURNAL PUBMED

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Insert length: 1771 Std Error: 0.00

Seq primer: T7

High quality sequence stop: 297.

FEATURES

source

1. .349
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3796098"
/db_xref="taxon:9606"
/clone="IMAGE:246852"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/notes="Organ: Liver and Spleen; Vector: p7T73D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(AT) primer
[5' ACTGGAAGATTAATTAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p7T73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 59.5%; Score 23.8; DB 2; Length 355;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATCTTGGCGGTAGAGCTGCTGGTCA 35
|||||
Db 314 ATCCAAATCTTGGCGGTAGGATGCTTTTCA 280

ORIGIN

Query Match 59.5%; Score 23.8; DB 8; Length 349;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATCTTGGCGGTAGAGCTGCTGGTCA 35
|||||
Db 169 ATCCAAATCTTGGCGGTAGGATGCTTTTCA 135

RESULT 9 BG952796/c

LOCUS BG952796 355 bp mRNA linear EST 12-JUN-2001
DEFINITION PM4-CT0543-200101-001-c06_1 CT0543 Homo sapiens cDNA, mRNA
sequence.

ACCESSION BG952796

VERSION BG952796.1 GI:14370967

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

1 (bases 1 to 355)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

PUBMED

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4&t2=PM4-CT0543-
200101-001-c06_1&t3=2001-01-20&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 355.

FEATURES

source

1. .355
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0543"
/notes="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORFESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN

Query Match 59.5%; Score 23.8; DB 2; Length 355;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATCTTGGCGGTAGAGCTGCTGGTCA 35
|||||
Db 314 ATCCAAATCTTGGCGGTAGGATGCTTTTCA 280

RESULT 10

BQ318970/c

LOCUS BQ318970 355 bp mRNA linear EST 17-MAY-2002

DEFINITION PM4-CT0543-290900-001-c06_1 CT0543 Homo sapiens cDNA, mRNA
sequence.

ACCESSION BQ318970

VERSION BQ318970.1 GI:20924739

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

1 (bases 1 to 355)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

10737800

JOURNAL

PUBMED

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4&t2=PM4-CT0543-
290900-001-c06_1&t3=2000-09-29&t4=1)

FEATURES	Seq primer: puc 18 forward	
	High quality sequence stop: 355.	
source	Location/Qualifiers	
	1..355	
ORIGIN	/organism="Homo sapiens"	
	/mol_type="mRNA"	
Query Match	59.5%; Score 23.8; DB 3; Length 355;	
	Best Local Similarity 80.0%; Pred. No. 1.3e+02;	
Matches	28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
	1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35	
Db		
	78 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 44	
RESULT 12	BF767919/c	
	LOCUS	
DEFINITION	367 bp mRNA linear EST 12-JAN-2001	
	CM1-CN0061-201200-673-g09 CN0061 Homo sapiens cDNA, mRNA sequence.	
ACCESSION	BF767919	
	VERSION	
KEYWORDS	BF767919.1 GI:12115819	
	EST.	
SOURCE	Homo sapiens (human)	
	ORGANISM	
REFERENCE	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
	Hominidae; Homo.	
TITLE	1 (bases 1 to 367)	
	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,	
JOURNAL	Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, P.F.,	
	Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,	
PUBMED	Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,	
	O'Hare, M.J., Soares, P., Brentani, R.R., Reis, L.F., de Souza, S.J. and	
COMMENT	Simpson, A.J.	
	Shotgun sequencing of the human transcriptome with ORF expressed	
FEATURES	sequence tags	
	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	
source	10737800	
	Contact: Simpson A.J.G.	
Embryos.	Laboratory of Cancer Genetics	
	Ludwig Institute for Cancer Research	
Query Match	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,	
	Brazil	
Best Local Similarity	Tel: +55-11-2704922	
	Fax: +55-11-2707001	
Conservative	Email: asimpson@ludwig.org.br	
	This sequence was derived from the FAPESP/LICR Human Cancer Genome	
Mismatches	Project. This entry can be seen in the following URL	
	(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM1&t2=CM1-CN0061-	
Indels	201200-673-g09&t3=2000-12-20&t4=1)	
	Seq primer: puc 18 forward	
Gaps	High quality sequence start: 3	
	High quality sequence stop: 364.	
Location/Qualifiers	Location/Qualifiers	
	1..367	
/organism="Homo sapiens"	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
/dev_stage="Adult"	/clone_lib="CN0061"	
	/note="Organ: colon normal; Vector: puc18; Site_1: SmaI;	
Vector: puc18; Site_1: SmaI;	Site_2: SmaI; A mini-library was made by cloning products	
	derived from ORESTES PCR (U.S. Letters Patent application	
No. 196,716 - Ludwig Institute for Cancer Research)	No. 196,716 - Ludwig Institute for Cancer Research)	
	profiles into the pUC 18 vector. Reverse transcription of	
tissue mRNA and cDNA amplification were performed under	low stringency conditions."	
	low stringency conditions."	
ORIGIN	Query Match	
	Best Local Similarity 80.0%; Pred. No. 1.3e+02;	
Matches	28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
	1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35	
Db		
	125 ATCCAAATTTCTGGCGGTAGAGAGCTGCTTTTCA 91	
Seq primer: puc 18 forward	High quality sequence stop: 355.	
	Location/Qualifiers	
1..355	/organism="Homo sapiens"	
	/mol_type="mRNA"	
59.5%; Score 23.8; DB 3; Length 355;	/db_xref="taxon:9606"	
	/dev_stage="Adult"	
80.0%; Pred. No. 1.3e+02;	/clone_lib="CT0543"	
	/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:	
0; Mismatches 7; Indels 0; Gaps 0;	SmaI; A mini-library was made by cloning products derived	
	from ORESTES PCR (U.S. Letters Patent application No.	
196,716 - Ludwig Institute for Cancer Research) profiles	196,716 - Ludwig Institute for Cancer Research) profiles	
	into the pUC 18 vector. Reverse transcription of tissue	
mRNA and cDNA amplification were performed under low	mRNA and cDNA amplification were performed under low	
	stringency conditions."	
stringency conditions."	59.5%; Score 23.8; DB 3; Length 355;	
	Best Local Similarity 80.0%; Pred. No. 1.3e+02;	
28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35	
358 bp mRNA linear EST 25-APR-2001	314 ATCCAAATTTCTGGCGGTAGAGAGCTGCTTTTCA 280	
358 bp mRNA linear EST 25-APR-2001	AW479483	
	25564 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	
ACCESSION	AW479483	
	VERSION	
KEYWORDS	AW479483.1 GI:7049589	
	EST.	
SOURCE	Bos taurus (cow)	
	ORGANISM	
REFERENCE	Bos taurus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;	
	Pecora; Bovidae; Bovinae; Bos.	
1 (bases 1 to 358)	Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,	
	Casas, E., Wray, J.B., White, J., Cho, J., Fahrenkrug, S.C.,	
Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,	Chitko-McKown, C.G., Perle, G., Holt, I., Karamycheva, S., Liang, F.,	
	Quackenbush, J. and Keele, J.W.	
Sequence evaluation of four pooled-tissue normalized bovine cDNA	libraries and construction of a gene index for cattle	
	Genome Res. 11 (4), 626-630 (2001)	
11282978	Contact: Smith TPL	
	USDA, ARS, US Meat Animal Research Center	
FO Box 166, Clay Center, NE 68933-0166, USA	Tel: 402 762 4366	
	Fax: 402 762 4390	
Email: smith@mail.marc.usda.gov	Single pass sequencing. Bases called and trimmed with phred	
	v0.980904.e. Vector identified by cross_match with the -minscore 20	
and -minmatch 12 options.	PCR Primers	
	FORWARD: AGGAACACGCTATGACCAT	
BACKWARD: GTTTCCTCAGTCACGACG	Plate: 12 row: K column: 12	
	Seq primer: ATTAGTGACACTATAG.	
Location/Qualifiers	Location/Qualifiers	
	1..358	
/organism="Bos taurus"	/mol_type="mRNA"	
	/db_xref="taxon:9913"	
/tissue_type="pooled"	/lab_host="DH108"	
	/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;	
Library made from pooled tissue from day 20 and day 40	Library made from pooled tissue from day 20 and day 40	

RESULT 13
AA159119/c
LOCUS
DEFINITION z059604.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone
INAGR:591198 5', mRNA sequence.

ACCESSION AA159119 436 bp mRNA linear EST 09-MAR-1998
VERSION z059604.r1
KEYWORDS INAGR:591198 5', mRNA sequence.

SOURCE
EST.
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

1 (bases 1 to 436)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, P.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 829 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 232.

FEATURES
source
1..436
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4622584"
/db_xref="taxon:9606"
/clone="IMAGE:591198"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene pancreas (#937208)"
/note="Organ: pancreas; Vector: pBluescript SK-; Site: 1:
ECORI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Pancreatic adenocarcinoma cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5'
CTCGAGCTTTTCTTTTCTTTT 3'"

ORIGIN
Query Match 59.5%; Score 23.8; DB 1; Length 436;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATCCAGATCTTCGGCGTAGAGAGCTGCTTGTC 35
|||||
DB 205 ATCCAAATTCCTGGCGTAGGATGCTTTTCA 171
|||||

RESULT 14
CD673609/c
LOCUS
DEFINITION CD673609 443 bp mRNA linear EST 24-JUN-2003
fs02e10.y1 Human Lens cDNA (Normalized): fs Homo sapiens cDNA clone
fs02e10 5', mRNA sequence.

ACCESSION CD673609
VERSION CD673609
KEYWORDS EST.
SOURCE CD673609.1 GI:32175340
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

1 (bases 1 to 443)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, P.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 829 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 232.

FEATURES
source
1..443
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4622584"
/db_xref="taxon:9606"
/clone="IMAGE:591198"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene pancreas (#937208)"
/note="Organ: pancreas; Vector: pBluescript SK-; Site: 1:
ECORI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Pancreatic adenocarcinoma cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5'
CTCGAGCTTTTCTTTTCTTTT 3'"

ORIGIN
Query Match 59.5%; Score 23.8; DB 1; Length 436;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATCCAGATCTTCGGCGTAGAGAGCTGCTTGTC 35
|||||
DB 205 ATCCAAATTCCTGGCGTAGGATGCTTTTCA 171
|||||

AUTHORS Wistow, G., Bernstein, S.L., Wyatt, M.K., Behal, A., Touchman, J.W.,
Bouffard, G., Smith, D. and Peterson, K.
Expressed sequence tag analysis of adult human lens for the NRIBank
Project: over 2000 non-redundant transcripts, novel genes and
splice variants
Mol. Vis. 8 (4), 171-184 (2002)
12107413
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 02 row: e column: 10
Seq primer: M3RP1 reverse primer (ABI).
Location/Qualifiers
1..443
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fs02e10"
/tissue type="Lens"
/dev stage="Adult"
/lab_host="EMDH108"
/clone_lib="Human Lens cDNA (Normalized): fs"
/note="Organ: Eye; Vector: pCMVSPORT6; A human lens
library (by) was normalized by self-subtraction. One
portion of double stranded plasmid DNA representing the
library was linearized by NotI. This NotI digested library
was used as a template for biotinylated RNA synthesis
using SP6 RNA polymerase. Another portion of the double
stranded plasmid library was converted to single-stranded
circles in vitro using Gene II and Exonuclease III (Life
Technologies). Single-stranded DNA (1 mg) was hybridized
(COT 500) with 41 mg of Bio-RNA and vector blocking
oligonucleotides. The hybridized Bio-RNA/ss-circles were
removed by streptavidin:phenol extraction. EST analysis
was performed on the library at the NIH Intramural
Sequencing Center (NISC)."

ORIGIN
Query Match 59.5%; Score 23.8; DB 6; Length 443;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATCCAGATCTTCGGCGTAGAGAGCTGCTTGTC 35
|||||
DB 164 ATCCAAATTCCTGGCGTAGGATGCTTTTCA 130
|||||

RESULT 15
BE485052/c
LOCUS
DEFINITION BE485052 471 bp mRNA linear EST 27-MAR-2003
171917 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE485052
VERSION BE485052.1 GI:9604585
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 471)
Sonstegard, T., Capuco, A.V., White, J., Van Tassel, C.P.,
Connor, E.E., Cho, J., Sultana, R., Shade, L., Wray, J.E., Wells, K.D.
and Quackenbush, J.
Analysis of bovine mammary gland EST and functional annotation of
the Bos taurus gene index
Mamm. Genome 13 (7), 373-379 (2002)
12140684
Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center

Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tads@lpsl.barc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 134 row: A column: 20
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1. .471
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="BARC 5BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."
FEATURES
source
Query Match 59.5%; Score 23.8; DB 2; Length 471;
Best Local Similarity 80.0%; Pred.No. 1.3e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATCCAGATTCTTGGCGGTAGAGAGCTGTGGTCA 35
|||||
Db 133 ATCCAGATTCTTGGCGGTGGGATGCTCTTCA 99
|||||

Search completed: April 18, 2006, 20:10:22
Job time : 1710.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:28:37 ; Search time 73.5 Seconds
(without alignments)
967.380 Million cell updates/sec

Title: SEQ4-THEN-SEQ3

Perfect score: 40

Sequence: 1 accagattcttggcgtagagctgttgcagactt 40

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 0.1

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/1_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/H_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 7: /cgn2_6/ptodata/1/ina/PP_COMB.seq.*
- 8: /cgn2_6/ptodata/1/ina/RE_COMB.seq.*
- 9: /cgn2_6/ptodata/1/ina/backfileseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	23.8	59.5	7827	3	US-09-620-312D-104
C 2	22.6	56.5	1600	3	US-09-938-540-1
C 3	21.8	54.5	663	3	US-09-248-796A-983
C 4	21.6	54.0	23193	3	US-09-949-016-17215
C 5	21	52.5	172677	3	US-09-949-016-13444
C 6	20.8	52.0	359	3	US-08-956-1718-4081
C 7	20.8	52.0	359	3	US-08-781-986A-4081
C 8	20.8	52.0	3191	3	US-09-270-767-13105
C 9	20.8	52.0	3786	3	US-08-961-527-182
C 10	20.6	51.5	35803	3	US-09-949-016-11863
C 11	20.6	51.5	35804	3	US-09-949-016-12962
C 12	20.4	51.0	629	3	US-09-477-135A-123
C 13	20.4	51.0	1627	3	US-09-270-767-28449
C 14	20.4	51.0	1891	3	US-09-270-767-12646
C 15	20.4	51.0	2602	3	US-09-902-540-405
C 16	20.4	51.0	4403765	3	US-09-103-840A-2
C 17	20.4	51.0	4411529	3	US-09-103-840A-1
C 18	20.2	50.5	1484	3	US-09-991-181-292
C 19	20.2	50.5	1484	3	US-09-990-444-292
C 20	20.2	50.5	1484	3	US-09-997-333-292
C 21	20.2	50.5	1484	3	US-09-992-598-292
C 22	20.2	50.5	767677	3	US-09-949-016-12147
C 23	20.2	50.5	767677	3	US-09-949-016-17361
C 24	20	50.0	20	3	US-09-938-642-3

25	20	50.0	20	3	US-09-938-642-4	Sequence 4, Appli
26	20	50.0	20	3	US-09-938-540-3	Sequence 3, Appli
27	20	50.0	20	3	US-09-938-540-4	Sequence 4, Appli
C 28	20	50.0	408	3	US-09-583-110-1764	Sequence 1764, Ap
C 29	20	50.0	408	3	US-09-107-433-2110	Sequence 2110, Ap
30	20	50.0	1600	3	US-09-938-540-1	Sequence 1, Appli
31	19.8	49.5	195	3	US-09-248-796A-9102	Sequence 9102, Ap
32	19.8	49.5	685	3	US-09-533-559-4705	Sequence 4705, Ap
C 33	19.8	49.5	1809	3	US-09-489-038A-4523	Sequence 4523, Ap
C 34	19.8	49.5	2244	3	US-09-949-016-4231	Sequence 4231, Ap
C 35	19.8	49.5	2230	3	US-09-620-312D-825	Sequence 825, App
C 36	19.8	49.5	21511	3	US-09-902-540-1201	Sequence 1201, Ap
C 37	19.8	49.5	37254	3	US-09-949-016-15973	Sequence 15973, A
38	19.6	49.0	601	3	US-09-949-016-96387	Sequence 96387, A
39	19.6	49.0	601	3	US-09-949-016-96388	Sequence 96388, A
40	19.6	49.0	601	3	US-09-949-016-96389	Sequence 96389, A
41	19.6	49.0	601	3	US-09-949-016-96653	Sequence 96653, A
42	19.6	49.0	601	3	US-09-949-016-96654	Sequence 96654, A
43	19.6	49.0	601	3	US-09-949-016-96655	Sequence 96655, A
44	19.6	49.0	601	3	US-09-949-016-96919	Sequence 96919, A
45	19.6	49.0	601	3	US-09-949-016-96920	Sequence 96920, A

ALIGNMENTS

RESULT 1
US-09-620-312D-104/c
; Sequence 104, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 656962el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt.FL_genes Version 1.0
; SEQ ID NO 104
; LENGTH: 7827
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(7497)
US-09-620-312D-104

Query Match 59.5%; Score 23.8; DB 3; Length 7827;
Best Local Similarity 80.0%; Pred. No. 5.9;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCGATTCTTGGCGGTAGAGAGCTGTGTGTC A 35

```
Db      6063 ATCCAAATCTTGGCGTAGGGATGCTTTTCA 6029
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 2
US-09-938-540-1/c
; Sequence 1, Application US/09938540
; Patent No. 6838267
; GENERAL INFORMATION:
; APPLICANT: Degussa AG
; TITLE OF INVENTION: New nucleotide sequences which code for the ccpA1 gene
; FILE REFERENCE: 000059 BT
; CURRENT APPLICATION NUMBER: US/09/938,540
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225)..(1388)
; OTHER INFORMATION: ccpA1-Gen
US-09-938-540-1

Query Match      56.5%; Score 22.6; DB 3; Length 1600;
Best Local Similarity 75.7%; Pred. No. 13;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 ATCCAGATCTTGGCGGTAGAGAGCTGCTTGGTCAGA 37
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      978 ATCCAGATCTTGGCGGTAGCCACCGTAATAGACAGA 942

RESULT 3
US-09-248-796A-983/c
; Sequence 983, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 983
; LENGTH: 663
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-983

Query Match      54.5%; Score 21.8; DB 3; Length 663;
Best Local Similarity 78.8%; Pred. No. 24;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1 ATCCAGATCTTGGCGGTAGAGAGCTGCTTGGT 33
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      140 ATCCAGTCTTGAAGGTAAGATATCTTGGT 108

RESULT 4
US-09-949-016-17215/c
; Sequence 17215, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13444
; LENGTH: 172677
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17215

Query Match      52.5%; Score 21; DB 3; Length 172677;
Best Local Similarity 82.8%; Pred. No. 1.5e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      9 TCTTGGCGGTAGAGAGCTGCTTGGTCAGA 37
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      10082 TCTTAGAGAGAGAGAGCTGCTTGGTCAGA 10054

RESULT 6
US-08-956-171E-4081/c
; Sequence 4081, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
```

```
;
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 4081:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4081:
US-08-956-171E-4081

Query Match 52.0%; Score 20.8; DB 3; Length 359;
Best Local Similarity 78.1%; Pred. No. 55;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 GATCTTGGCGGTAGAGAGCTGCTTGGTCAGA 37
Db 263 GATTATTGGAGGTAGAGCACTGTTGGTACGA 232

RESULT 7
US-08-781-986A-4081/c
; Sequence 4081, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
```

```
;
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 4081:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-4081

Query Match 52.0%; Score 20.8; DB 3; Length 359;
Best Local Similarity 78.1%; Pred. No. 55;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 GATCTTGGCGGTAGAGAGCTGCTTGGTCAGA 37
Db 263 GATTATTGGAGGTAGAGCACTGTTGGTACGA 232

RESULT 8
US-09-270-767-13105/c
; Sequence 13105, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13105
; LENGTH: 3191
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; US-09-270-767-13105

Query Match 52.0%; Score 20.8; DB 3; Length 3191;
Best Local Similarity 78.1%; Pred. No. 83;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 ATTCTTGGCGGTAGAGAGCTGCTTGGTCAGAC 38
Db 109 ATTGTGGGGGAGAGAGAGCTGCTGGGCTAGAC 78

RESULT 9
US-08-961-527-182
; Sequence 182, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
```


Search completed: April 18, 2006, 19:31:14
Job time : 76.5 secs

```
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28449
; LENGTH: 1627
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-28449

Query Match      51.0%; Score 20.4; DB 3; Length 1627;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTT 30
   ||| ||||| ||||| ||||| ||||| |||||
DB 793 ACCATATTCTTGTCTTAGAGAGCTGCCT 822

RESULT 14
US-09-270-767-12646
; Sequence 12646, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12646
; LENGTH: 1891
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-12646

Query Match      51.0%; Score 20.4; DB 3; Length 1891;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTT 30
   ||| ||||| ||||| ||||| ||||| |||||
DB 1057 ACCATATTCTTGTCTTAGAGAGCTGCCT 1086

RESULT 15
US-09-902-540-405/c
; Sequence 405, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 405
; LENGTH: 2602
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-405

Query Match      51.0%; Score 20.4; DB 3; Length 2602;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 AGATTCCTTGGCGGTAGAGAGCTGCTGTC 34
   ||| ||||| ||||| ||||| ||||| |||||
DB 1487 AGAGTCTTGGAGGTAGAGCACTGATTGGAC 1458
```

This Page Blank (uspto)


```

; FILE REFERENCE: UTSD:773US
; CURRENT APPLICATION NUMBER: US/10/080,381B
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 60/291,354
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/274,706
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/270,251
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 1640
; TYPE: DNA
; ORGANISM: RAT
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (4)..(1628)
; OTHER INFORMATION: N = A, C, G, T or U
US-10-080-381B-67

```

```

Query Match          61.5%; Score 24.6; DB 6; Length 1640;
Best Local Similarity 76.9%; Pred. No. 2.2;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

QY      1 ATCCGATTCTTGGCGGTAGAGAGCTGCTTGGTGCACT 39
          |||||
Db      1277 ATCCGATTCTTGGCGGTAGACACAGCTTGCCCACT 1315

```

```

RESULT 3
US-09-822-846-554/c
; Sequence 554, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulkota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; CURRENT APPLICATION NUMBER: US/09/822,846
; PRIOR FILING DATE: 2001-03-29
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 554
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-846-554

```

```

Query Match          59.5%; Score 23.8; DB 3; Length 860;
Best Local Similarity 80.0%; Pred. No. 4.4;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY      1 ATCCGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35

```

```

Db      55 ATCCAAATCTTGGCGGTAGGATGCTTTTCA 21
          |||||

```

```

RESULT 4
US-09-822-846-570/c
; Sequence 570, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulkota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6400
; CURRENT APPLICATION NUMBER: US/09/822,846
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,605
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 570
; LENGTH: 1928
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-846-570

```

```

Query Match          59.5%; Score 23.8; DB 3; Length 1928;
Best Local Similarity 80.0%; Pred. No. 4.9;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY      1 ATCCGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35
          |||||
Db      176 ATCCAAATCTTGGCGGTAGGATGCTTTTCA 142
          |||||

```

```

RESULT 5
US-10-264-049-616/c
; Sequence 616, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 616
; LENGTH: 3134
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

```
; LOCATION: (319)..(319)
; OTHER INFORMATION: n equals a.t.g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3108)..(3108)
; OTHER INFORMATION: n equals a.t.g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3117)..(3117)
; OTHER INFORMATION: n equals a.t.g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3132)..(3132)
; OTHER INFORMATION: n equals a.t.g, or c
; US-10-264-049-616

Query Match          59.5%; Score 23.8; DB 6; Length 3134;
Best Local Similarity 80.0%; Pred. No. 5.2;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1330 ATCCAAATTCTTGGCGGTAGGATGCTTTTCA 1296

RESULT 6
US-10-291-172-69/c
; Sequence 69, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 69
; LENGTH: 7747
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (41)..(7414)
; US-10-291-172-69

Query Match          59.5%; Score 23.8; DB 6; Length 7747;
Best Local Similarity 80.0%; Pred. No. 5.9;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5983 ATCCAAATTCTTGGCGGTAGGATGCTTTTCA 5949

RESULT 7
US-10-221-278-69/c
; Sequence 69, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
```

```
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 69
; LENGTH: 7747
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (41)..(7414)
; US-10-221-278-69

Query Match          59.5%; Score 23.8; DB 7; Length 7747;
Best Local Similarity 80.0%; Pred. No. 5.9;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5983 ATCCAAATTCTTGGCGGTAGGATGCTTTTCA 5949

RESULT 8
US-10-119-926-21/c
; Sequence 21, Application US/10119926
; Publication No. US20030104413A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Felyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Xu, Chongjun
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104413A1el Nucleic Acids and
; FILE REFERENCE: 789CIP2BCON
; CURRENT APPLICATION NUMBER: US/10/119,926
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 21
; LENGTH: 7827
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(7497)
; US-10-119-926-21

Query Match          59.5%; Score 23.8; DB 5; Length 7827;
```

```
Best Local Similarity 80.0%; Pred. No. 5.9;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGCTCA 35
Db 6063 ATCCAAATTCTTGGCGCTAGGGATGCTTTTCA 6029

RESULT 9
US-10-037-270-104/c
; Sequence 104, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 104_
; LENGTH: 7827
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(7497)
US-10-037-270-104

Query Match 59.5%; Score 23.8; DB 5; Length 7827;
Best Local Similarity 80.0%; Pred. No. 5.9;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGCTCA 35
Db 6063 ATCCAAATTCTTGGCGCTAGGGATGCTTTTCA 6029

RESULT 10
US-10-117-722-104/c
; Sequence 104, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/117,722

QY 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGCTCA 35
Db 6063 ATCCAAATTCTTGGCGCTAGGGATGCTTTTCA 6029

RESULT 11
US-10-122-851-104/c
; Sequence 104, Application US/10122851
; Publication No. US20050239060A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/122,851
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 104_
; LENGTH: 7827
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(7497)
US-10-122-851-104

Query Match 59.5%; Score 23.8; DB 9; Length 7827;
Best Local Similarity 80.0%; Pred. No. 5.9;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGCTCA 35
Db 6063 ATCCAAATTCTTGGCGCTAGGGATGCTTTTCA 6029

RESULT 12
US-10-210-172-35/c
; Sequence 35, Application US/10210172
; Publication No. US20040043928A1
; GENERAL INFORMATION:
```

APPLICANT: Kekuda, Ramesh
APPLICANT: Miller, Charles
APPLICANT: Patturajan, Meera
APPLICANT: Pena, Carol
APPLICANT: Rieger, Daniel
APPLICANT: Shimkets, Richard
APPLICANT: Zerhusen, Bryan
APPLICANT: Li, Li
APPLICANT: Ji, Weizhen
APPLICANT: Padigaru, Muralidhara
APPLICANT: Casman, Stacie
APPLICANT: Voss, Edward
APPLICANT: Boldog, Ferenc
APPLICANT: Gorman, Linda
APPLICANT: Leite, Mario
APPLICANT: Vernet, Corine
APPLICANT: Anderson, David
APPLICANT: Guo, Xiaojia
APPLICANT: Zhong, Mei
APPLICANT: Gerlach, Valerie
APPLICANT: Hjalt, Tord
APPLICANT: Rastelli, Luca
APPLICANT: Spytek, Kimberly
APPLICANT: Edinger, Shlomit
APPLICANT: Ellerman, Karen
APPLICANT: Malyankar, Uriel
APPLICANT: MacDougall, John
APPLICANT: Stone, David
APPLICANT: Alsobrook II, John
APPLICANT: Lepley, Denise et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-416 A
CURRENT APPLICATION NUMBER: US/10/210,172
CURRENT FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/309,501
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 60/323,994
PRIOR FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: 60/373,814
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/310,291
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 60/310,951
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: 60/310,544
PRIOR FILING DATE: 2001-08-07
PRIOR APPLICATION NUMBER: 60/311,292
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/311,979
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 60/313,201
PRIOR FILING DATE: 2001-08-17
PRIOR APPLICATION NUMBER: 60/312,892
PRIOR FILING DATE: 2001-08-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 327
SOFTWARE: Curaseqlist version 0.1
SEQ ID NO 35
LENGTH: 9508
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (170)..(9187)
US-10-210-172-35

Query Match 59.5%; Score 23.8; DB 7; Length 9508;
Best Local Similarity 80.0%; Pred. No. 6;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35
Db 7756 ATCCAAATTTCTTGGCGGTAGAGGTGCTTTTCA 7722

RESULT 13
US-10-369-493-24140/c
Sequence 24140, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 24140
LENGTH: 2157
TYPE: DNA
ORGANISM: Xenorhabdus nematophilus
US-10-369-493-24140

Query Match 59.0%; Score 23.6; DB 6; Length 2157;
Best Local Similarity 76.3%; Pred. No. 6.1;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 38
Db 399 ATCCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 362

RESULT 14
US-10-501-282-995/c
Sequence 995, Application US/10501282
Publication No. US20050203280A1
GENERAL INFORMATION:
APPLICANT: MCMICHAEL, JOHN CALHOUN
APPLICANT: ZAGURSKY, ROBERT JOHN
APPLICANT: RUSSELL, DAVID PARRISH
APPLICANT: FLETCHER, LEAH DIANE
TITLE OF INVENTION: POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
FILE REFERENCE: AM100780 L2
CURRENT APPLICATION NUMBER: US/10/501,282
CURRENT FILING DATE: 2004-07-09
PRIOR APPLICATION NUMBER: 60/333,777
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 60/426,742
PRIOR FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: PCT/US02/36123
PRIOR FILING DATE: 2002-11-25
NUMBER OF SEQ ID NOS: 6653
SOFTWARE: PatentIn version 3.2
SEQ ID NO 995
LENGTH: 1266
TYPE: DNA
ORGANISM: Alloiococcus otitidis
FEATURE:
NAME/KEY: CDS
LOCATION: (121)..(1263)
US-10-501-282-995

Query Match 57.5%; Score 23; DB 9; Length 1266;
Best Local Similarity 83.9%; Pred. No. 10;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 AGATTCTTGGCGGTAGAGAGCTGCTTGGTCA 35
Db 915 AAATTGTTGGAGGTAGCTAGCTGCTTGGTCA 885

```
RESULT 15
US-10-501-282-997/c
; Sequence 997, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; PRIOR FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 997
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Alloiococcus otitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(1263)
US-10-501-282-997

Query Match      57.5%; Score 23; DB 9; Length 1266;
Best Local Similarity 83.9%; Pred. No. 10;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      5  AGATTCTTGGCGGTAGAGAGCTGTGGTCA 35
Db      915 AAATTGTGGAGGTAGCTAGCTGTGGTCA 885

Search completed: April 18, 2006, 23:13:57
Job time : 680.5 secs
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 19:49:40 ; Search time 352 Seconds
(without alignments)
457.708 Million cell updates/sec

Title: SEQ4-THEN-SEQ3
Perfect score: 40
Sequence: 1 atccagattcttggcggtagagagctgcttgtagactt 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 9281099 seqs, 2013915447 residues

Total number of hits satisfying chosen parameters: 18562198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.New.*

- 1: /SIDSS/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 2: /SIDSS/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /SIDSS/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /SIDSS/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 5: /SIDSS/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 6: /SIDSS/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 7: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 8: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 9: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 10: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 11: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 12: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 13: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 14: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 15: /SIDSS/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22.2	55.5	1706	14	US-11-136-527-2411
C 2	21.4	53.5	663	6	US-09-925-065A-918089
C 3	20.6	51.5	518	6	US-09-925-065A-591215
C 4	20.6	51.5	535	6	US-09-925-065A-591216
C 5	20.6	51.5	535	9	US-10-301-480-10681
C 6	20.6	51.5	535	10	US-10-301-480-624090
C 7	20.6	51.5	557	10	US-10-301-480-594607
C 8	20.6	51.5	557	10	US-10-301-480-1208016
C 9	20.6	51.5	557	10	US-10-301-480-478124
C 10	20.6	51.5	557	10	US-10-301-480-1091533
C 11	20.6	51.5	596	6	US-09-925-065A-411990
C 12	20.6	51.5	1381	8	US-10-750-185-48123
C 13	20.6	51.5	1381	8	US-10-750-623-48123
C 14	20.2	50.5	60	6	US-09-925-065A-757534
C 15	20.2	50.5	1073	14	US-11-000-463-507
C 16	20.2	50.5	1073	14	US-11-000-463-508
C 17	20.2	50.5	1379	14	US-11-000-463-36
C 18	20.2	50.5	1484	7	US-10-194-487-219

C 19	20.2	50.5	1484	7	US-10-195-883-219	Sequence 219, App
C 20	20.2	50.5	1484	7	US-10-195-888-219	Sequence 219, App
C 21	20.2	50.5	1484	7	US-10-195-889-219	Sequence 219, App
C 22	20.2	50.5	1508	14	US-11-124-367A-194	Sequence 194, App
C 23	20	50.0	436	6	US-09-925-065A-508353	Sequence 508353, App
C 24	20	50.0	2190	8	US-10-750-185-26101	Sequence 26101, A
C 25	20	50.0	2190	8	US-10-750-623-26101	Sequence 26101, A
C 26	20	50.0	2645	8	US-10-750-185-43843	Sequence 43843, A
C 27	20	50.0	2645	8	US-10-750-623-43843	Sequence 43843, A
C 28	19.8	49.5	92	8	US-10-310-914A-16251	Sequence 16251, A
C 29	19.8	49.5	201	8	US-10-995-561-69138	Sequence 69138, A
C 30	19.8	49.5	201	8	US-10-995-561-69145	Sequence 69145, A
C 31	19.8	49.5	567	6	US-09-925-065A-608235	Sequence 608235, App
C 32	19.8	49.5	567	6	US-09-925-065A-608236	Sequence 608236, App
C 33	19.8	49.5	576	10	US-10-301-480-427081	Sequence 427081, App
C 34	19.8	49.5	576	10	US-10-301-480-1040490	Sequence 1040490, App
C 35	19.8	49.5	581	6	US-09-925-065A-355591	Sequence 355591, App
C 36	19.8	49.5	640	6	US-09-925-065A-697463	Sequence 697463, App
C 37	19.8	49.5	704	11	US-11-031-206-165	Sequence 165, App
C 38	19.8	49.5	1059	8	US-10-750-185-34935	Sequence 34935, A
C 39	19.8	49.5	1059	8	US-10-750-623-34935	Sequence 34935, A
C 40	19.8	49.5	2837	6	US-09-925-065A-690684	Sequence 690684, App
C 41	19.8	49.5	3202	14	US-11-120-308-189	Sequence 189, App
C 42	19.8	49.5	3236	11	US-11-031-206-167	Sequence 167, App
C 43	19.8	49.5	23803	8	US-10-995-561-13406	Sequence 13406, A
C 44	19.6	49.0	543	6	US-09-925-065A-25010	Sequence 25010, A
C 45	19.6	49.0	543	9	US-10-301-480-126247	Sequence 126247, App

ALIGNMENTS

RESULT 1

US-11-136-527-2411/c
; Sequence 2411, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2411
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2411

Query Match 55.5%; Score 22.2; DB 14; Length 1706;
Best Local Similarity 77.1%; Pred. No. 22;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 CCAGATTCTTGGCGGTAGAGAGCTGCTGCTGCTGAGA 37
DB 428 CCAGATTGGGGCGGTGGAGAGCTGCTGCTGCTGAGA 394

RESULT 2

US-09-925-065A-918089
; Sequence 918089, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08

```
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 918089
/ LENGTH: 663
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-918089

Query Match      53.5%; Score 21.4; DB 6; Length 663;
Best Local Similarity 80.6%; Pred. No. 41;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 9 TCTTGGCGGTAGAGAGCTGCTTGGTCAGACT 39
Db 540 TATTGTGGTAGAGAGCTGCTTGGTCAGGCT 570

RESULT 3
US-09-925-065A-591215
/ Sequence 591215, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 591215
/ LENGTH: 518
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-591215

Query Match      51.5%; Score 20.6; DB 6; Length 518;
Best Local Similarity 74.3%; Pred. No. 87;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 6 GATTCCTGGCGGTAGAGAGCTGCTTGGTCAGACTT 40
Db 235 GATGTGGGTGGTAAAGAGCTCCCTAGACAGACTT 269

RESULT 4
US-09-925-065A-591216
/ Sequence 591216, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
```

```
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 591216
/ LENGTH: 535
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-591216

Query Match      51.5%; Score 20.6; DB 6; Length 535;
Best Local Similarity 74.3%; Pred. No. 87;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 6 GATTCCTGGCGGTAGAGAGCTGCTTGGTCAGACTT 40
Db 245 GATGTGGGTGGTAAAGAGCTCCCTAGACAGACTT 279

RESULT 5
US-10-301-480-10681
/ Sequence 10681, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10681
/ LENGTH: 535
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-10681

Query Match      51.5%; Score 20.6; DB 9; Length 535;
Best Local Similarity 74.3%; Pred. No. 87;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 6 GATTCCTGGCGGTAGAGAGCTGCTTGGTCAGACTT 40
Db 245 GATGTGGGTGGTAAAGAGCTCCCTAGACAGACTT 279

RESULT 6
US-10-301-480-624090
/ Sequence 624090, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
```

```
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 624090
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-624090
```

```
Query Match 51.5%; Score 20.6; DB 10; Length 535;
Best Local Similarity 74.3%; Pred. No. 87;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 6 GATCTTGGCGGTAGAGAGCTGCTTGGTCAGACTT 40
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 GATGTGGGTGTAAGAGCTCCTAGACAGACTT 279
```

```
RESULT 7
US-10-301-480-594607
; Sequence 594607, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 594607
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-594607
```

```
Query Match 51.5%; Score 20.6; DB 10; Length 557;
Best Local Similarity 74.3%; Pred. No. 88;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 6 GATCTTGGCGGTAGAGAGCTGCTTGGTCAGACTT 40
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 GATTATTGGTTGCAGAGCTCCTGTTGTAACCTT 201
```

```
RESULT 8
US-10-301-480-1208016
; Sequence 1208016, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1208016
; LENGTH: 557
; TYPE: DNA
```

```
; ORGANISM: Homo sapien
US-10-301-480-1208016
```

```
Query Match 51.5%; Score 20.6; DB 10; Length 557;
Best Local Similarity 74.3%; Pred. No. 88;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 6 GATCTTGGCGGTAGAGAGCTGCTTGGTCAGACTT 40
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 GATTATTGGTTGCAGAGCTCCTGTTGTAACCTT 201
```

```
RESULT 9
US-10-301-480-478124/c
; Sequence 478124, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 478124
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-478124
```

```
Query Match 51.5%; Score 20.6; DB 10; Length 592;
Best Local Similarity 74.3%; Pred. No. 89;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 6 GATCTTGGCGGTAGAGAGCTGCTTGGTCAGACTT 40
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 49 GATAATTGCAGGAAGAGCACTTGTGTCAGACTT 15
```

```
RESULT 10
US-10-301-480-1091533/c
; Sequence 1091533, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1091533
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1091533
```

```
Query Match 51.5%; Score 20.6; DB 10; Length 592;
Best Local Similarity 74.3%; Pred. No. 89;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 6 GATCTTGGCGGTAGAGAGCTGCTTGGTCAGACTT 40
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 49 GATAATTGCGAGGAGCAACTTTGTCAGACTT 15
RESULT 11
US-09-925-065A-411990
; Sequence 411990, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 411990
; LENGTH: 596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-411990
Query Match 51.5%; Score 20.6; DB 6; Length 596;
Best Local Similarity 74.3%; Pred. No. 89; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 9;
QY 6 GATCTTGGCGGTAGAGAGCTGCTTGGTCAGACTT 40
Db 544 GATAATTGCGAGGAGCAACTTTGTCAGACTT 578
RESULT 12
US-10-750-185-48123/c
; Sequence 48123, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48123
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Bovine 19866881441129
US-10-750-185-48123
Query Match 51.5%; Score 20.6; DB 8; Length 1381;
Best Local Similarity 74.3%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 9;
QY 2 TCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCAG 36
Db 1276 TACAGAAGTTTGGAGGTAGGTAGCTGCTGGTTTCAG 1242
RESULT 13
US-10-750-623-48123/c
; Sequence 48123, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48123
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Bovine 19866881441129
US-10-750-623-48123
Query Match 51.5%; Score 20.6; DB 8; Length 1381;
Best Local Similarity 74.3%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 9;
QY 2 TCAGATTCTTGGCGGTAGAGAGCTGCTTGGTCAG 36
Db 1276 TACAGAAGTTTGGAGGTAGGTAGCTGCTGGTTTCAG 1242
RESULT 14
US-09-925-065A-757534/c
; Sequence 757534, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 757534
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-757534
Query Match 50.5%; Score 20.2; DB 6; Length 640;
Best Local Similarity 75.8%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 8;
QY 5 AGATTCTTGGCGGTAGAGAGCTGCTTGGTCAGA 37
```

Db 583 AGATGTTAGCAGGAGAGAGCTGCTGGGCAGA 551

RESULT 15
US-11-000-463-507/c
; Sequence 507, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 507
; LENGTH: 1073
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-000-463-507

Query Match 50.5%; Score 20.2; DB 14; Length 1073;
Best Local Similarity 75.8%; Pred. No. 1.5e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 5 AGATCTTGGCGGTAGAGAGCTGCTGGTCAGA 37
| | | | | | | | | | | | | | | | | | | | | |
Db 691 AGATGTTAGCAGGAGAGAGCTGCTGGGCAGA 659

Search completed: April 18, 2006, 20:01:33
Job time : 353 secs

This Page Blank (uspto)